

R;Stover, C.K.; Pham, X.Q.; Ervin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: A83253
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-316 <STO>
A/Cross-references: GB:AE004738; GB:AE004091; NID:G949252; PIDN:AA606534.1; GSPDB:GN001
A/Experimental source: strain PA01
C/Genetics:
A/Gene: wbpK; PA3146

Query Match 68.5%; Score 37; DB 2; Length 316;
Best Local Similarity 72.7%; Pred. No. 44;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 12
:|||||
DB 24 VIVVRRVRR 34

RESULT 8
P98146
probable amino-acid ABC transporter permease protein y4tP AGR_L_252 [imported] - Agrobac
C/Species: Agrobacterium tumefaciens
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C/Accession: P98146
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
; Liu, P.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A/Reference number: A97359; PMID:11743194
A/Accession: P98146
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-221 <KUR>
A/Cross-references: GB:AE007870; PIDN:AAK8696.1; PID:G15158429; GSPDB:GN00170
C/Genetics:
A/Gene: AGR_L_252
A/Map position: linear chromosome
C/Superfamily: histidine permease protein M

Query Match 66.7%; Score 36; DB 2; Length 221;
Best Local Similarity 54.5%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 12
:|||||
DB 201 IIRVIEHVRR 211

RESULT 9
AF3141
hypothetical protein Atu4754 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C/Accession: AF3141
R;Wood, D.W.; Sebudal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavln, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; PMID:11743193
A/Accession: AF3141
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-221 <KUR>
A/Cross-references: GB:AE008669; PIDN:AAL45548.1; PID:G17743262; GSPDB:GN00187

A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu4754
A/Map position: linear chromosome
C/Superfamily: histidine permease protein M

Query Match 66.7%; Score 36; DB 2; Length 221;
Best Local Similarity 54.5%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 12
:|||||
DB 201 IIRVIEHVRR 211

RESULT 10
T29709
probable aspartate transaminase (EC 2.6.1.1) C14E2.2 [similarity] - *Caenorhabditis eleg*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C/Accession: T29709
R;Bentley, D.; Gattung, S.
Submitted to the EMBL Data Library, January 1996
A/Description: The sequence of C. elegans cosmid C14E2.
A/Reference number: Z20670
A/Accession: T29709
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-364 <BEN>
A/Cross-references: EMBL:U46671; PIDN:AA85747.1; CESP:C14E2.2
C/Genetics:
A/Gene: CESP:C14E2.2
A/Introns: 89/2; 213/1; 316/1
C/Superfamily: aspartate aminotransferase
C/Keywords: aminotransferase

Query Match 66.7%; Score 36; DB 2; Length 364;
Best Local Similarity 72.7%; Pred. No. 74;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVVRVRRVRR 11
:|||||
DB 74 REVVARRVRR 84

RESULT 11
E70602
probable membraneprotein weakly - *Mycobacterium tuberculosis* (strain H37RV)
C/Species: *Mycobacterium tuberculosis*
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C/Accession: E70602
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A/Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: E70602
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-503 <COL>
A/Cross-references: GB:Z94752; GB:AL123456; NID:G3261731; PIDN:CB08157.1; PID:el299780;
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: RV1002c

Query Match 66.7%; Score 36; DB 2; Length 503;
Best Local Similarity 70.0%; Pred. No. 16+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 11
:|||||

L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebut, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
C:Contents: annotation
C:Genetics:
A:Gene: Sma556
A:Genome: plasmid

Query Match 72.2%; Score 39; DB 2; Length 561;
Best Local Similarity 72.7%; Pred. No. 35;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRRVRRVRR 12
DB 208 VLRVRRRLRR 218

RESULT 3

AF3371
dihydropteroate synthase (EC 2.5.1.15) [imported] - *Brucella melitensis* (strain 16M)

C:Species: *Brucella melitensis*
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AF3371

R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-446, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3371

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-279 <KIR>

A:Cross-references: GB:AE008917; PIDN:AAU52137.1; PID:G17982913; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME10956

A:Map position: 1

C:Superfamily: dihydropteroate synthase; dihydropteroate synthase homology
C:Keywords: transferase

Query Match 70.4%; Score 38; DB 2; Length 279;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVRRVRRVRR 12
DB 81 RVRRVRRVRR 218

RESULT 4

H72708
probable ribosomal protein L31 APE1087 - *Aeropyrum pernix* (strain K1)

C:Species: *Aeropyrum pernix*

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 02-Aug-2002

C:Accession: H72708

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamaya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72708

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-105 <KAW>

A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAAG0072.1; PID:G1043858; PID:G510

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1087

C:Superfamily: rat ribosomal protein L31

Query Match 68.5%; Score 37; DB 2; Length 105;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVRRVRRVRR 12
DB 28 RAVRRVRRVRR 39

RESULT 5

UT0588
hypothetical 20K protein (celly region) - *Erwinia chrysanthemi*

C:Species: *Erwinia chrysanthemi*

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Sep-1999

C:Accession: UT0588

R:Guiseppi, A.; Aymeric, J.L.; Cami, B.; Barras, F.; Crezzer, N.

Gene 106, 109-114, 1991

A:Title: Sequence analysis of the cellulase-encoding celly gene of *Erwinia chrysanthemi*:
A:Reference number: UT0585; MUID:92039050; PMID:1937031

A:Accession: UT0588

A:Molecule type: DNA

A:Residues: 1-170 <GUI>

A:Cross-references: GB:M74044

A:Experimental source: strain 3937

C:Superfamily: *Erwinia chrysanthemi* hypothetical 20K protein (celly region)

Query Match 68.5%; Score 37; DB 2; Length 170;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVRRVRRVRR 12
DB 147 RAVRRVRRVRR 158

RESULT 6

C75508
hypothetical protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000

C:Accession: C75508

R:White, O.; Eiben, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; W
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: C75508

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-276 <WHI>

A:Cross-references: GB:AE001911; GB:AE000513; NID:G6458217; PIDN:AAFI0115.1; PID:G64582

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0533

A:Map position: 1

C:Superfamily: *Deinococcus radiodurans* hypothetical protein DR0533
Query Match 68.5%; Score 37; DB 2; Length 276;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVRRVRRVRR 12
DB 229 RAVRRVRRVRR 240

RESULT 7

A83253
probable NAD-dependent epimerase/dehydratase WbpK PA3146 [imported] - *Pseudomonas aerug*:
C:Species: *Pseudomonas aeruginosa*
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: A83253

C:Superfamily: rat ribosomal protein L31

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:48:06 ; Search time 4.97872 Seconds

(without alignments)
231.709 Million cell updates/sec

Title: US-10-079-075-4

Perfect score: 54

Sequence: 1 RVRVRVRRVRR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:***
2: PIR:***
3: PIR:***
4: PIR:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	79.6	2351	2 G71415	hypothetical protein
2	39	72.2	561	2 G95367	probable methyl-ace
3	38	70.4	279	2 AF3371	dihydropterolate sy
4	37	68.5	105	2 H72708	probable ribosomal
5	37	68.5	170	2 J70588	hypothetical 20K p
6	37	68.5	276	2 C75508	hypothetical prote
7	37	68.5	316	2 A83253	probable NAD-depen
8	36	66.7	221	2 F98146	probable amino-act
9	36	66.7	221	2 AF3141	hypothetical prote
10	36	66.7	354	2 T29709	probable aspartate
11	36	66.7	503	2 E70602	probable membrane
12	36	66.7	740	2 S43566	R01H10.8 protein (
13	36	66.7	1677	2 T14267	kin protein, stage
14	35	64.8	87	2 I40094	hypothetical prote
15	35	64.8	117	2 A75220	hypothetical prote
16	35	64.8	125	2 F87326	ISCC3, transposase
17	35	64.8	247	2 T17677	hypothetical prote
18	35	64.8	510	2 H86932	probable membrane
19	35	64.8	732	2 C84487	hypothetical prote
20	34	63.0	164	2 B70922	hypothetical prote
21	34	63.0	185	2 G83223	conserved hypotet
22	34	63.0	196	1 H5RT19	ribosomal protein
23	34	63.0	196	2 A46554	ribosomal protein
24	34	63.0	196	2 A48992	uridine kinase udk
25	34	63.0	229	2 G83809	hypothetical prote
26	34	63.0	229	2 T35717	hypothetical prote
27	34	63.0	250	2 T35717	conserved hypotet
28	34	63.0	310	2 AF0335	conserved hypotet
29	34	63.0	310	2 AG0804	conserved hypotet

30	34	63.0	319	2 AE0759	Chib protein limpo
31	34	63.0	421	2 F91030	probable adenine-s
32	34	63.0	421	2 G85874	probable adenine-s
33	34	63.0	421	2 H65005	site-specific DNA-
34	34	63.0	613	2 G69429	Glu-tRNA amidotran
35	34	63.0	648	2 C69423	DNA helicase homol
36	34	63.0	972	2 AB3640	isolectine-cRNA 11
37	34	63.0	1082	2 H81982	hypothetical prote
38	34	63.0	1231	2 C84716	hypothetical prote
39	34	63.0	2049	2 T29227	hypothetical prote
40	33	61.1	32	2 I51089	proteamine - Japane
41	33	61.1	121	2 E87679	chemotaxis protein
42	33	61.1	135	2 T11919	ribosomal protein
43	33	61.1	173	2 AH3486	soluble lytic mure
44	33	61.1	258	2 T36264	probable dehydroge
45	33	61.1	314	2 H82115	probable adenine-s

ALIGNMENTS

RESULT 1

G71415

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C>Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998

C/Accession: G71415

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Ditt, P.; Wedler, H.; Wedler, E.; Wammut, R.; Weizenecker, T.; Pohl, T.M.; Terry, N.; Gle, avenagh, T.; Hempel, S.; Kotter, P.; Entlan, K.D.; Rieger, M.; Schaeffer, W.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulidomenei, erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Anr

C: Chalvatzis, N.

A>Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: G71415

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-2351 <REV>

A:Cross-references: GB:I297338; NID:92244870; PID:e327462; PID:92244876

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match 79.6%; Score 43; DB 2; Length 2351;

Best Local Similarity 58.3%; Pred. No. 28;

Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVRVRVRRVRR 12

Db 69 RIVKIVRKVKR 80

RESULT 2

G95367

probable methyl-accepting chemotaxis protein [imported] - Sinorhizobium meliloti (tetrair

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C/Accession: G95367

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe

.; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: G95367

A:Status: preliminary

A:Molecule type: DNA

A:Cross-references: GB:AE006469; PIND:AAK5505.1; PID:g14523978; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

peita, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.

Mon Jun 9 12:26:39 2003

us-10-079-075-12.ra1

Page 7

```

Oy 19 EVFVTVRRVRRV--RRVRRV-----YFVV--RRVFV 48
      | : | | | | | : | | | | |
Db 26 RISRVCRRWRRLVDDRWLMRRHVDLTLYTRALAGRAWAAV 65

```

Search completed: June 9, 2003, 12:05:10
Job time : 16.3404 secs


```

? TELEPHONE: 617-345-9100
? TELEFAX: 617-945-9111
? INFORMATION FOR SEQ ID NO: 27:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 31 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: peptide
? SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-150-812-37

```

Query Match	21.4%	Score 53.5;	DB 4;	Length 31;
Best Local Similarity	44.8%	Pred. No. 3.8;		
Matches 13; Conservative	0;	Mismatches 11;	Indels 5;	Gaps 1;

```

QY      14 RRVWRVVRRVRRRV-----RRVRRVWR 37
          || ||| | | | | | | | | |
Db      2  RRAWRRAKRRRAARCGVSARRAARRAWR 30

```

```

: RESULT 11
: US-09-413-814-48
: Sequence 48, Application US/09413814
: Patent No. 6225064
: GENERAL INFORMATION:
: APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
: APPLICANT: Bristol-Myers Squibb, Co.
: APPLICANT: Beyer, Stefan
: APPLICANT: Bloecker, Helmut
: APPLICANT: Brandt, Petra
: APPLICANT: Cino, Paul M
: APPLICANT: Dougherty, Brian A
: APPLICANT: Goldberg, Steven L
: APPLICANT: Hofle, Gerhard
: APPLICANT: Mueller, Joachim
: TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
: TITLE OF INVENTION: heteropolysaccharide compounds
: FILE REFERENCE: PCT/US 99/23535
: CURRENT APPLICATION NUMBER: US/09/413,814
: CURRENT FILING DATE: 1999-10-07
: EARLIER APPLICATION NUMBER: DE 198 46 493.2
: EARLIER FILING DATE: 1998-10-09
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 48
: LENGTH: 2475
: TYPE: PRT
: ORGANISM: Sorangium cellulosum
: US-09-413-814-48

```

Query Match	21.2%	Score 53	DB 4	Length 2475
Best Local Similarity	39.4%	Pred No. 2.4e+02		
Matches 13, Conservative	4	Mismatches 16	Indels 0	Gaps 0

Qy 6 VRRWVRVRRVWRVRRVRRVRRVRRVRRV 38
 |||| : : ||| |||
 Db 1802 VRRWQGLGELWRAGMPVWRERQGRGRHAGRRV 1834

RESULT 12
US-09-199-637A-289
Sequence 289, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Austbel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Mehajan-Miklos, Shalina
APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenkeard, Eliana

```

1  APPLICANT: Tsongalis, John
2  TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
3  TITLE OF INVENTION: SEQUENCES AND USES THEREOF
4  FILE REFERENCE: 007867361002
5  CURRENT APPLICATION NUMBER: US/09/199,637A
6  CURRENT FILING DATE: 1998-11-25
7  PRIOR APPLICATION NUMBER: 60/066,517
8  PRIOR FILING DATE: 1997-11-25
9  NUMBER OF SEQ ID NOS: 437
10 SOFTWARE: FastSeq for Windows Version 4.0
11 SEQ ID NO 289
12 LENGTH: 101
13 TYPE: PR1
14 ORGANISM: Pseudomonas aeruginosa
15 US-09-199-637A-289

```

Query Match	20.6%	Score 51.5;	DB 4;	Length 101;
Best Local Similarity	42.5%	Pred. No. 18;		
Matches 17; Conservative	0;	Mismatches 18;	Indels 5;	Gaps 2

```
OY      11 RRVRRVRRVVRRVR---VRRVMR-RVRRVRRW 45
          |  |  |  |  |  |  |  |  |  |  |
Db      41 RARRSSRRSSRPSRRWRRGARATSRAMRASSTRPSRTW 80
```

RESULT 13
 US-08-286-819A-28
 Sequence 28, Application US/08286819A
 Patent No. 5871910
 GENERAL INFORMATION:
 APPLICANT: ARTHUR, MICHEL
 APPLICANT: DUKTA-MALEN, SYLVIE
 APPLICANT: MOLINAS, CATHERINE
 APPLICANT: COURVALIN, PATRICE
 TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
 TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
 TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
 NUMBER OF SEQUENCES: 54
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/286,819A
 FILING DATE: 05-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/174,682
 FILING DATE: 28-DEC-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/917,146
 FILING DATE: 10-AUG-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/FR/91/00855
 FILING DATE: 29-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 9013579
 FILING DATE: 31-OCT-1990
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Oblon, NO. 5871910man F.

```

: REFERENCE/DOCKET NUMBER: 234R2-PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (515) 248-4500
: TELEFAX: (515) 248-4844
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
PCT-US95-00062-9

```

Query Match	23.4%	Score 58.5	DB 5	length 31
Best Local Similarity	41.9%	Pred. No. 1.1		
Matches 13, Conservative	6	Mismatches 9	Indels 3	Gaps 1

```
Qy      11 RRVRRVVRVVVVRRVRRVRRVRRVRRV 41
        ||:| |:| :||: ||: ||
Db       1 RRIYRAIRHIIPIRIGW--LRIIGRIEEV 28
```

RESULT 8
US-08-995-172-1
; Sequence 1, Application US/08995172B

```

: APPLICANT: Thatcher, David R
: APPLICANT: Milks, Paula E
: TITLE OF INVENTION: Optimization of Gene Delivery and Gene Delivery Systems
: FILE REFERENCE: CAC000026
: CURRENT APPLICATION NUMBER: US/08/995,172B
: CURRENT FILING DATE: 1997-12-22
: EARLIER APPLICATION NUMBER: 60/033,908
: EARLIER FILING DATE: 1996-12-23
: NUMBER OF SEQ. ID NOS.: 25
: SOFTWARE: PatentIn Ver. 2.1
: SEQ. ID NO. 1
: LENGTH: 31

```

OTHER INFORMATION: Description of Artificial Sequence: Peptide
US-08-995-172-1

Query Match	21.4%	Score	53.5	DB	4	Length	31
Best Local Similarity	44.8%	Pred.No.	3.8				
Matches	13	Conservative	0	Mismatches	11	Indels	5
						Gaps	1

Qy 14 RRVRRVVRVVRVWV-----RVVRVWR 37
||| |
Db 2 RRAWRRAKRRAARRCGVSARARAARRR 30

RESULT 9
US-08-839-624-27
; Sequence 27, Application US/08839624

GENERAL INFORMATION:
APPLICANT: Kain et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
TITLE OF INVENTION: HIV INFECTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:

```
STREET: One Financial Center  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk
```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,624
FILING DATE: April 15, 1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/78191
FILING DATE: 15-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5350
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown

MOLECULE TYPE: peptide
IS-08-839-624-27

```

Query March	21.4%	Score 53.5	DB 4	Length 31
Best Local Similarity	44.8%	Pred. No. 3	8	
Matches 13, Conservative	0	Mismatches	11	Indels 5; Gaps 1

```

OY      14 RRWRVRVVRVVRV-----RRVRVWR 37
          ||||| |
Db       2 RRWRRAKRRARRCGVSARRARRWR 30

```

RESULT 10
US-09-150-812-27
; Sequence 27, Application US/09150812

APPLICANT: Karn et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATTING HIV INFECTION

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Banner & Witcoff, Inc
;; STREET: One Financial Center
;; CITY: Boston

STATE: Massachusetts
COUNTRY: USA
ZIP: 02111

```
;
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: WordPerfect 6.1
```

```

; FILING DATE: 11-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/839,624

```

1 APPLICATION NUMBER: US 60/017,268
2 FILING DATE: 13-MAY-1996
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Kathleen M. Williams
5 REGISTRATION NUMBER: 34,380
6 REFERENCE/DOCKET NUMBER: 3255/5399
7 TELECOMMUNICATION INFORMATION:

TITLE OF INVENTION: WEIGHT HEPAPIN
 TITLE OF INVENTION: WEIGHT HEPAPIN
 TITLE OF INVENTION: ANTICOAGULATION REVERSAL
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Benita J. Rohm, Esq.
 STREET: 6601 Woodward Avenue
 STREET: Suite 1525
 CITY: Detroit
 STATE: Michigan
 COUNTRY: United States of America
 ZIP: 48226
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk 1.44mb, 3.5"
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: Wordperfect 6;
 SOFTWARE: ASCII (DOS)Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/436,703B
 FILING DATE: 08-May-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: N/A
 FILING DATE: N/A
 ATTORNEY/AGENT INFORMATION:
 NAME: Rohm, Benita J.
 REGISTRATION NUMBER: 28,664
 REFERENCE/DOCKET NUMBER: 7WK-060548-00233
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 313-965-1976
 TELEFAX: 313-965-1951
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 38 amino acids
 TYPE: amino acid
 STRANDEDNESS: N/A
 TOPOLOGY: N/A
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: N/A
 PUBLICATION INFORMATION:
 AUTHORS: N/A
 TITLE: N/A
 JS-08-436-703B-17

Query Match 24.4%; Score 61; DB 2; Length 38;
Best Local Similarity 48.6%; Pred. No. 0.7;
Matches 18; Conservative 1; Mismatches 14; Indels 4; Gaps 1;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
82

COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: wordperfect 6/
SOFTWARE: ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7WK-060548-00233
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-5

Query Match	24.0%;	Score 60;	DB 2;	Length 39;
Best Local Similarity	51.7%;	Pred. No. 0.92;		
Matches 15; Conservative	0;	Mismatches 14;	Indels 0;	Gaps 0;

Oy 18 RRVVRRVRVRVRVRVRVRVRVRVRVR 46
| | | | | | | |
Db 3 RRAARRRARRARRARRARRARRAR 31

```

/ RESULT 4
/ US-09-413-814--78
/ Sequence 78, Application US/09413814
/ Patent No. 6225064
/ GENERAL INFORMATION:
/ APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
/ APPLICANT: Bristol-Myers Squibb, Co.
/ APPLICANT: Beyer, Stefan
/ APPLICANT: Bloecker, Helmut
/ APPLICANT: Brandt, Petra
/ APPLICANT: Cino, Paul M
/ APPLICANT: Dougherty, Brian A
/ APPLICANT: Goldberg, Steven L
/ APPLICANT: Hofte, Gerhard
/ APPLICANT: Mueller, Joachim
/ APPLICANT: Reichenbach, Hans
/ TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
/ TITLE OF INVENTION: heteropolypeptide compounds
/ FILE REFERENCE: PCT/US 99/23535
/ CURRENT APPLICATION NUMBER: US/09/413,814
/ CURRENT FILING DATE: 1999-10-07
/ EARLIER APPLICATION NUMBER: DE 198 46 493.2
/ EARLIER FILING DATE: 1998-10-09
/ NUMBER OF SEQ ID NOS: 107
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 78
/ LENGTH: 882
/ TYPE: PRT

```


OTHER INFORMATION: Clone ID: UC-ZMFLMO17108D04_FLI.pep
US-10-425-114-68542

Query Match 30.0%; Score 75; DB 6; Length 209;
Best Local Similarity 41.3%; Pred. No. 10;
Matches 26; Conservative 3; Mismatches 8; Indels 26; Gaps 4;

QY 4 RVRVR-----VRRVRVVRVVRVVRVVR-----VRRVRVVRVVR 43
DB 78 RGVRRGRGLGRVCGRLGAMSVRRRGRGLGRR--RLGRRGVRRRGRGLGCRRLMRGV----- 131
QY 44 RWR 46
DB 132 RWR 134

RESULT 7
US-10-425-114-71061
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 71061
LENGTH: 252
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLMO17030B05_FLI.pep
US-10-425-114-71061

Query Match 30.0%; Score 75; DB 6; Length 252;
Best Local Similarity 41.3%; Pred. No. 12;
Matches 26; Conservative 3; Mismatches 8; Indels 26; Gaps 4;

QY 4 RVRVR-----VRRVRVVRVVRVVRVVR-----VRRVRVVRVVR 43
DB 78 RGVRRGRGLGRVCGRLGAMSVRRRGRGLGRR--RLGRRGVRRRGRGLGCRRLMRGV----- 131
QY 44 RWR 46
DB 132 RWR 134

RESULT 8
US-10-425-114-68594
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 68594
LENGTH: 342
TYPE: PRT
ORGANISM: Zea mays

FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLMO17159C10_FLI.pep
US-10-425-114-68594

Query Match 30.0%; Score 75; DB 6; Length 342;
Best Local Similarity 41.3%; Pred. No. 15;
Matches 26; Conservative 3; Mismatches 8; Indels 26; Gaps 4;

QY 4 RVRVR-----VRRVRVVRVVRVVRVVR-----VRRVRVVRVVR 43
DB 78 RGVRRGRGLGRVCGRLGAMSVRRRGRGLGRR--RLGRRGVRRRGRGLGCRRLMRGV----- 131
QY 44 RWR 46
DB 132 RWR 134

RESULT 9
US-10-425-114-58367
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 58367
LENGTH: 129
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3591-091-B12_FLI.pep
US-10-425-114-58367

Query Match 28.4%; Score 71; DB 6; Length 129;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 18; Conservative 1; Mismatches 6; Indels 8; Gaps 2;

QY 8 RVRVRVVRVVRVVRVVR-----VRRVRVVRVVR 38
DB 12 RVRRLRRRR-----RRVVRVVRVVRVVRVVR 38

RESULT 10
US-60-452-680-16830
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: GRUPE, Andrew
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CL001450
CURRENT APPLICATION NUMBER: US/60/452,680
CURRENT FILING DATE: 2003-03-07
NUMBER OF SEQ ID NOS: 116213
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16830
LENGTH: 142
TYPE: PRT
ORGANISM: Homo sapiens
US-60-452-680-16830

Query Match 28.4%; Score 71; DB 7; Length 142;
Best Local Similarity 38.6%; Pred. No. 17;
Matches 22; Conservative 3; Mismatches 20; Indels 12; Gaps 4;

```

? SEQ ID NO 66861
? LENGTH: 197
? TYPE: PRT
? ORGANISM: Zea mays
? FEATURES
? OTHER INFORMATION: Clone ID: UC-2MFUBJ320IH12_FL1.pptf
? US-10-425-114-66861

```

Query Match	30.2%	Score 75.5	DB 6	Length 197
Best Local Similarity	53.7%	Pred. No. 8.9		
Matches 22; Conservative		1; Mismatches	11; Indels	7; Gaps 4

```

QY      8 RWRVRVRVRVRVRVRVRVR--VVRVRVRVRVRVRVRVR 46
          |||||  |||  |||  |||  |||  |||  |||
DB      58 RWRPRCRWRWTR--RWTATWTCWRIRTPWR--VR-SRWR 93

```

```

RESULT 3
US-10-264-237-2376
; Sequence 2376, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL31P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2376
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (90)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-2376

```

		Query Match	30.0%; Score 75; DB 6;	length 136;
D8		Best Local Similarity	30.0%;	Pred. No. 7, 4'
Matches	18;	Conservative	14;	Mismatches 12; Indels 16; Gaps 3
CY	2 VVVVVR-----RVRR- :::	RVMR-RVTVRVVRVRVRVRVRVRVR :::		45
D8	18 ILRLRNWLMLLLLRNRLRLRLILLWMNRLLITLWRRLRLMLILTRRW :::			77

RESULT 4
 US-10-425-114-70810
 Sequence 70810, Application US/10425114
 GENERAL INFORMATION:
 APPLICANT: Liu, Jingtong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack B
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5313) B
 CURRENT APPLICATION NUMBER: US/10/425, 114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 70810
 LENGTH: 190
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: UC-ZMFM017240G10_PL1 pep

US-10-425-114-70810

Query Match	30.0%	Score 75;	DB 6;	Length 190;
Best Local Similarity	41.3%;	Pred. No. 9.6;		
Matches	26;	Conservative	3;	Mismatches 8;
				Indels 26;
				Gaps 4;

```
QY      4 RVRWRW-----VRRVRWRVRVRVRWRVR----VRRWRVRVRV. 43
          |||||         ||| : ||| : |||||   ||:|||
DB     78 RGVRWRGRLGRVCGRLGAMSVRRRGLGRR--RLRGVRVRRRGRLGGCRLMWKGV--- 133
```

QY	44	RWR	46
Db	132	RWR	134

```

RESULT 5
US-10-425-114-68513
: Sequence 68513, Application US/10425114 .
: GENERAL INFORMATION:
: APPLICANT: Liu, Jindong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovalic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(5313)B
: CURRENT APPLICATION NUMBER: US/10/425.114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 68513
: LENGTH: 195
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: UC-ZMFLMO17297F07_FLI pep
US-10-425-114-68513

```

```

Query Match      30.0%; Score 75; DB 6; Length 19;
Best Local Similarity 41.3%; Pred. No. 9.8;
Matches          26; Conservative    3; Mismatches    8; Indels   26; Gaps  4;

Qy      4 RVRRRW-----VRRRVRRRVVVRRRVR-----VRRVRRRVVRRV 43
        |||||               ||| : ||| ||| ||| ||| |||
Db       78 KGVRRWGRLGRVCGLGAGSVRRRRGRIGRR--FLGGRCVRRRGRLGCCRLMRKGV--- 131
        |||||               ||| : ||| ||| ||| ||| |||

Qy      44 RWR 46
        |||
Db       132 RWR 134
```

```

RESULT 6
US-10-425-114-68542
: Sequence 68542, Application US/10425114
: GENERAL INFORMATION:
: APPLICANT: Liu, Jindong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovalic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(5313)B
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 68542
: LENGTH: 209
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:

```



```
; Sequence 8, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-8
```

```
Query Match      68.0%; Score 170; DB 1; Length 48;
Best Local Similarity 90.9%; Pred. No. 6,7e-12;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 RRVVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 44
Db 1 RRVVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 44
```

```
RESULT 13
US-09-785-058-8
```

```
; Sequence 8, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-8
```

```
Query Match      68.0%; Score 170; DB 21; Length 48;
Best Local Similarity 90.9%; Pred. No. 6,7e-12;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 RRVVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 44
Db 1 RRVVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 44
```

```
RESULT 14
US-09-785-059-8
```

```
; Sequence 8, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A3577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
```

```
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-8
```

```
Query Match      68.0%; Score 170; DB 21; Length 48;
Best Local Similarity 90.9%; Pred. No. 6,7e-12;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 RRVVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 44
Db 1 RRVVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 44
```

```
RESULT 15
```

```
US-10-079-075-8
; Sequence 8, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-8
```

```
Query Match      68.0%; Score 170; DB 24; Length 48;
Best Local Similarity 90.9%; Pred. No. 6,7e-12;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 RRVVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 44
Db 1 RRVVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 44
```

```
Search completed: June 9, 2003, 12:25:30
Job time : 208.83 secs
```

[illegible]

```
RESULT 2
PCT-US02-04812-12
; Sequence 12, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
PCT-US02-04812-12
```

```
Query Match          100.0%; Score 250; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 6.4e-21; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RVVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV
Db 1 RVVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 48
```

```
RESULT 3
US-09-785-058-12
; Sequence 12, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-058-12
```

```
Query Match          100.0%; Score 250; DB 21; Length 48;
Best Local Similarity 100.0%; Pred. No. 6.4e-21; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RVVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV
Db 1 RVVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 48
```

```
RESULT 4
US-09-785-059-12
; Sequence 12, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
```

```
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-059-12
```

```
Query Match          100.0%; Score 250; DB 21; Length 48;
Best Local Similarity 100.0%; Pred. No. 6.4e-21; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RVVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV
Db 1 RVVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 48
```

```
RESULT 5
US-10-079-075-12
; Sequence 12, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-10-079-075-12
```

```
Query Match          100.0%; Score 250; DB 24; Length 48;
Best Local Similarity 100.0%; Pred. No. 6.4e-21; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RVVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV
Db 1 RVVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 48
```

```
RESULT 6
PCT-US02-04432-11
; Sequence 11, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04432-11
```

```
Query Match          70.4%; Score 176; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e-12; Indels 0; Gaps 0;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 13 VRVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV
Db 1 VRVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 33
```

```
RESULT 7
```



```

; RESULT 14
; US-10-079-075-10
; Sequence 10, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: prt
; ORGANISM: Artificial Sequence

```

Search completed: June 9, 2003, 12:34:11
Job time : 25.5319 secs

XX	Otc C,	Komurian-Pradel F;
PI		
XX	WPI; 2002-049351/06.	
DR	N-PSDB; ABA04890.	
XX		
PT	New immunogenic polypeptide from TT virus, useful for diagnosis and as	
PT	therapeutic vectors, also related nucleic acid and antibodies -	
PS	Example 2; Fig 1; 45dp; French.	
XX		
CC	The present invention relates to a novel immunogenic protein from TT	
CC	virus (AA47794). TT virus is associated with non-A to G hepatitis.	
CC	The protein of the invention is useful as an immunogen useful for	
CC	detecting anti-TT virus antibodies and for raising specific antibodies	
CC	for detection of TT virus proteins. Antibodies raised against the protein	
CC	of the invention provide rapid and reliable identification of all TT	
CC	virus infections, regardless of genotype and/or subtype. The present	
CC	invention was used in an example from the present invention.	
XX		
SQ	Sequence	756 AA;
Query Match	30.8%;	Score 77; DB 23; Length 756;
Best Local Similarity	32.8%;	Pred. No. 0.39; *
Matches	22; Conservative	5; Mismatches 12; Indels 28; Gaps 3
OY	7 RRMVRRVVRRVVRRVVVR-----RMRVRRVVRRVVVR-----	40
Db	11 RRMRRRRRPMPRRRRRTTRRRRRPARRGGRGRRVVRRRRRGRRRRRR--MKKGRGRKKAK	69
	41 -VRRRR 46	
OY	:::::	
Db	70 IIRRWQ 76	
RESULT 7		
AAW9082		
ID	AAW9082 standard; Protein; 770 AA.	
XX		
AC	AAW9082;	
XX		
DT	13-MAY-1999 (first entry)	
XX		
DE	Non-B, non-C, non-G hepatitis virus protein sequence SEQ ID NO:9.	
XX		
KM	Hepatitis virus; non-B non-C non-G hepatitis virus; PCR primer;	
KW	virion; detection; screening; infection; blood transfusion.	
XX		
OS	Hepatitis virus.	
XX		
FN	MO9905282-A1.	
XX		
PD	04-FEB-1999.	
XX		
PF	27-JUL-1998; 98WO-JP03340.	
XX		
ER	13-MAR-1998; 98JP-0082962.	
PR	25-JUL-1997; 97JP-0233246.	
PR	09-OCT-1997; 97JP-0314196.	
XX		
PA	(TAMU/) TAMURA R.	
XX		
PI	Nishizawa T, Okamoto H;	
XX		
DR	WPI; 1999-142937/12.	
XX		
PT	Gene isolated from non-B non-C non-G hepatitis DNA virus - and its	
PT	expression products, useful for diagnosis and treatment of hepatitis	
PT	and screening of blood for transfection.	
XX		
FS	Claim 20; Page 84-87; 113pp; Japanese.	
XX		
CC	The present invention describes a gene which is contained in a non-B	

```
CC non-C non-G hepatitis DNA virus. It is 3500-4000 bases in length and contains two overlapping open reading frames (ORF). It is obtained by polymerase chain reaction (PCR) amplification using primer. The present sequence represents a specifically claimed non-B non-C non-G hepatitis protein sequence. The gene can be used for the production of vaccines for prevention and treatment of non-B non-C non-G hepatitis infection. CC Diagnosis of such infection, and screening of blood (e.g. intended for transfusion) for the presence of the virus, by using the virion or antigenic peptides as reagents for detection of antibodies to the virus, or by direct detection of the gene using PCR with primers derived from the gene sequence.
```

```
CC  
CC Sequence      770 AA;  
SQ
```

```
Query March           30.8%; Score 77; DB 20; Length 770;  
Best Local Similarity 32.8%; Pred. No. 0.4;  
Matches   22; Conservative    5; Mismatches   12; Indels   28; Gaps   3;
```

```
OY       7 RRVRRVRVRRVVRRVVVR-----RVRVRVRVRRVVVR----- 40  
Db        ||||| | | | |  
          11 RRVRRVRVRRPPRRRRRRRRRRRRPARRRGRRRRVRRRRGGWRRRRYRR-WRKGRRRKKAK 69  
OY         41 -VRRRMK 46  
          :::::  
Db          70 IIRQWQ 76
```

```
RESULT 8  
ID AAY97179 standard; Protein: 770 AA.  
XX AAY97179;  
AC  
DT 04-DEC-2000 (first entry)  
XX  
DE TT virus (TTV-GH1) open reading frame 1.  
XX  
KW TTV; TT virus; blood transmission; detection; amplification; primer;  
KM transplantation; xenotransplantation; vector; ORF1.  
XX  
OS TT virus.  
FH Key Location/Qualifiers  
FT Misc-difference 1..100 /label= Arginine_rich_motif  
FT FT Misc-difference 125..127 /label= Conserved_motif_1  
FT FT Active-site 126  
XX  
EN WO200046407-A2.  
XX  
PD 10-AUG-2000.  
XX  
PF 04-FEB-2000; 2000WO-US02982.  
XX  
PR 05-FEB-1999; 99US-0245248.  
XX  
PA (ABBO ) ABBOTT LAB.  
XX  
PI Leary TP, Simons JN, Erker JC, Chalmers ML, Birkenmeyer LG;  
PI Muernhoff AS, Pilot-Matias TJ, Desai SM, Mushahwar IK,  
XX WPI; 2000-514969/46.  
DR N-PSDB; AAA53632.  
XX  
PT New oligomer primer useful for the detection of TT virus in test  
PT samples and tissues and organs for use in (xeno)transplantation  
XX  
ES Example 3; Page 105; 139pp; English.  
XX
```

The TT virus (TTV) genome was obtained from the serum of a West African individual (GH1). It is a circular, negative single-stranded DNA virus. The GH1 isolate was found to be 3852 nucleotides in length. 113

Db :::|:|:
70 IIRQWQ 76

RESULT 4
AAB03840 ID AAB03840 standard; protein; 120 AA.
XX
AC AAB03840;
XX
DT 20-OCT-2000 (first entry)
XX
DE Protein fragment #2 used in TT virus antibody determination.
XX
KM Serum type classification; TT virus; antibody; viral infection;
XX identify; treatment.
OS TT virus.
XX JP2000135087-A.
XX
PD 16-MAY-2000.
XX
PF 29-OCT-1998; 98UP-0309208.
XX
PR 29-OCT-1998; 98JP-0309208.
XX
PA (SRLS-) SRL KK.
XX
DR WPI; 2000-415430/36.
XX
PT Peptides for determination of anti-TT virus antibody and method for
XX serum classification of TT virus using the peptides -
XX
PS Claim 1; Page 7; 12pp; Japanese.
XX
CC A method for serum type classification of TT virus (also known as
XX hepatitis TT virus) has been identified. The method relies on the use of
XX peptide fragments of the virus. The invention also relates to the use of
XX CC TT virus peptides for anti-TT virus antibody determination. The anti-TT
XX virus antibodies and the serum type classification method, can be used to
XX screen TT virus, to determine its route of infection, and seroconversion.
XX CC The classification of TT virus may lead to improved treatment of viral
XX disease. The present sequence represents a fragment of TT virus protein
XX used in the course of the invention.
SQ Sequence 120 AA;
Query Match 31.6%; Score 79; DB 21; length 120;
Best Local Similarity 55.0%; Pred. No. 0.037;
Matches 22; Conservative 1; Mismatches 15; Indels 2; Gaps 1;
OY 9 WRRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 48
6 WRRRRRRPWR--RRRRRRRRRRRRRRRRVRRYRRRTV 43
DB
RESULT 5
AAG28909 ID AAG28909 standard; Protein; 82 AA.
XX
AC AAG28909;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 34302.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PP 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135153.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139751.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140053.
XX 23-JUN-1999; 99US-0140054.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140623.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.

[illegible]

```

RX SEQUENCE FROM N.A.
RX MEDLINE=21844401; PubMed=11855633;
RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,
RA Okamoto H.;
RT "Analysis of the complete genomes of thirteen TT virus variants
RT classifiable into the fourth and fifth genetic groups, isolated from
RT viremic infants.";
RL Arch. Virol. 147:21-41(2002).
DR EMBL; AB064628; BAB9400.1; -.
FT NON TER
SQ SEQUENCE 49 AA; 7107 MW; DF3110365A07BE66 CRC64;

Query March 31.4%; Score 78.5; DB 12; Length 49;
Best Local Similarity 43.8%; Pred. No. 0.026; Mismatches 9; Indels 15; Gaps 3;
Matches 21; Conservative 3;

QY 9 WVRERVRVRVRVRVRVRVRVRVRVRVRVRV-----RVVRVRVRVR 46
DQ | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DQ 5 WKKKKRRRMMWR-----KKMTRGRRLRRMPRRSRRRPRRRRRRRRRR 47

RESULT 12
O71097 PRELIMINARY; PRT; 171 AA.
ID O71097
AC O71097;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE PVII.
OS Bovine adenovirus type 3 (Mastadenovirus bos3).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
CX NCBI_TaxID=10510;
N1
N2
N3
N4
N5
N6
N7
N8
N9
N10
N11
N12
N13
N14
N15
N16
N17
N18
N19
N20
N21
N22
N23
N24
N25
N26
N27
N28
N29
N30
N31
N32
N33
N34
N35
N36
N37
N38
N39
N40
N41
N42
N43
N44
N45
N46
N47
N48
N49
N50
N51
N52
N53
N54
N55
N56
N57
N58
N59
N60
N61
N62
N63
N64
N65
N66
N67
N68
N69
N70
N71
N72
N73
N74
N75
N76
N77
N78
N79
N80
N81
N82
N83
N84
N85
N86
N87
N88
N89
N90
N91
N92
N93
N94
N95
N96
N97
N98
N99
N100
N101
N102
N103
N104
N105
N106
N107
N108
N109
N110
N111
N112
N113
N114
N115
N116
N117
N118
N119
N120
N121
N122
N123
N124
N125
N126
N127
N128
N129
N130
N131
N132
N133
N134
N135
N136
N137
N138
N139
N140
N141
N142
N143
N144
N145
N146
N147
N148
N149
N150
N151
N152
N153
N154
N155
N156
N157
N158
N159
N160
N161
N162
N163
N164
N165
N166
N167
N168
N169
N170
N171
N172
N173
N174
N175
N176
N177
N178
N179
N180
N181
N182
N183
N184
N185
N186
N187
N188
N189
N190
N191
N192
N193
N194
N195
N196
N197
N198
N199
N200
N201
N202
N203
N204
N205
N206
N207
N208
N209
N210
N211
N212
N213
N214
N215
N216
N217
N218
N219
N220
N221
N222
N223
N224
N225
N226
N227
N228
N229
N230
N231
N232
N233
N234
N235
N236
N237
N238
N239
N240
N241
N242
N243
N244
N245
N246
N247
N248
N249
N250
N251
N252
N253
N254
N255
N256
N257
N258
N259
N260
N261
N262
N263
N264
N265
N266
N267
N268
N269
N270
N271
N272
N273
N274
N275
N276
N277
N278
N279
N280
N281
N282
N283
N284
N285
N286
N287
N288
N289
N290
N291
N292
N293
N294
N295
N296
N297
N298
N299
N300
N301
N302
N303
N304
N305
N306
N307
N308
N309
N310
N311
N312
N313
N314
N315
N316
N317
N318
N319
N320
N321
N322
N323
N324
N325
N326
N327
N328
N329
N330
N331
N332
N333
N334
N335
N336
N337
N338
N339
N340
N341
N342
N343
N344
N345
N346
N347
N348
N349
N350
N351
N352
N353
N354
N355
N356
N357
N358
N359
N360
N361
N362
N363
N364
N365
N366
N367
N368
N369
N370
N371
N372
N373
N374
N375
N376
N377
N378
N379
N380
N381
N382
N383
N384
N385
N386
N387
N388
N389
N390
N391
N392
N393
N394
N395
N396
N397
N398
N399
N400
N401
N402
N403
N404
N405
N406
N407
N408
N409
N410
N411
N412
N413
N414
N415
N416
N417
N418
N419
N420
N421
N422
N423
N424
N425
N426
N427
N428
N429
N430
N431
N432
N433
N434
N435
N436
N437
N438
N439
N440
N441
N442
N443
N444
N445
N446
N447
N448
N449
N450
N451
N452
N453
N454
N455
N456
N457
N458
N459
N460
N461
N462
N463
N464
N465
N466
N467
N468
N469
N470
N471
N472
N473
N474
N475
N476
N477
N478
N479
N480
N481
N482
N483
N484
N485
N486
N487
N488
N489
N490
N491
N492
N493
N494
N495
N496
N497
N498
N499
N500
N501
N502
N503
N504
N505
N506
N507
N508
N509
N510
N511
N512
N513
N514
N515
N516
N517
N518
N519
N520
N521
N522
N523
N524
N525
N526
N527
N528
N529
N530
N531
N532
N533
N534
N535
N536
N537
N538
N539
N540
N541
N542
N543
N544
N545
N546
N547
N548
N549
N550
N551
N552
N553
N554
N555
N556
N557
N558
N559
N560
N561
N562
N563
N564
N565
N566
N567
N568
N569
N570
N571
N572
N573
N574
N575
N576
N577
N578
N579
N580
N581
N582
N583
N584
N585
N586
N587
N588
N589
N590
N591
N592
N593
N594
N595
N596
N597
N598
N599
N600
N601
N602
N603
N604
N605
N606
N607
N608
N609
N610
N611
N612
N613
N614
N615
N616
N617
N618
N619
N620
N621
N622
N623
N624
N625
N626
N627
N628
N629
N630
N631
N632
N633
N634
N635
N636
N637
N638
N639
N640
N641
N642
N643
N644
N645
N646
N647
N648
N649
N650
N651
N652
N653
N654
N655
N656
N657
N658
N659
N660
N661
N662
N663
N664
N665
N666
N667
N668
N669
N670
N671
N672
N673
N674
N675
N676
N677
N678
N679
N680
N681
N682
N683
N684
N685
N686
N687
N688
N689
N690
N691
N692
N693
N694
N695
N696
N697
N698
N699
N700
N701
N702
N703
N704
N705
N706
N707
N708
N709
N710
N711
N712
N713
N714
N715
N716
N717
N718
N719
N720
N721
N722
N723
N724
N725
N726
N727
N728
N729
N730
N731
N732
N733
N734
```

```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=PMV.
RX      MEDLINE=20409089; PubMed=10950985;
RA      Hallett R.L., Clewley J.P., Bobet F., McKiernan P.J., Teo C.G.;
RT      "Characterization of a highly divergent TT virus genome.";
RL      J. Gen. Virol. 81:2273-2279(2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=PMV.
RA      Hallett R.L., Clewley J.P., Teo C.G.;
RL      Submitted (Apr-2000) to the EMBL/genBank/DBJ databases.
DR      EMBL; AF261761; AAF8255.1; -
DR      InterPro: IPR004219: TTVirus_unk.
DR      pfam; PF02956; TT_ORF1; 1.
SQ      SEQUENCE 766 AA; 90844 MW; 22B5A3BA3CC8187 CRC64;

Query Match          32.6%; Score 81.5; DB 12; Length 766;
Best Local Similarity 42.6%; Pred. No. 0.15;
Matches    23; Conservative    1; Mismatches   15; Indels   15; Gaps    2.

Qy      7 RRWRVRVRVRVVVVV-----VRRWRVRVRVRVRVRVR 46
Db      11 RRPFRRRRTTTRRPPRRRRRTATRRRGVRVR-RRGGRGMRRTYTRRRR 63
ID      Q8V7C1 PRELIMINARY; PRT; 49 AA.
AC      Q8V7C1;
DT      01-MAR-2002 (TREMBLrel. 20, Created)
DT      01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT      01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE      ORF1 (fragment) .
OS      TT virus.
OC      Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX      NCBI_TaxID=68887;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21844401; PubMed=11855633;
RA      Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,
RA      Okamoto H.;
RT      "Analysis of the complete genomes of thirteen TT virus variants
RT      classifiable into the fourth and fifth genetic groups, isolated from
RT      viremic infants."
RL      Arch. Virol. 147:21-41(2002) .
DR      EMBL; AB064629; BAB79402.1; -.
FT      NON TER 49
SQ      SEQUENCE 49 AA; 7034 MW; C1EAGEBF6A1DCFCDC CRC64;

Query Match          32.2%; Score 80.5; DB 12; Length 49;
Best Local Similarity 45.8%; Pred. No. 0.015;
Matches    22; Conservative    2; Mismatches   9; Indels   15; Gaps    3;

Qy      9 WVRVRVRVRVRVRVRVR-VRRRV-----RRVVRVVRVR 46
Db      5 WKRRRRGGMMR-----RRWTGRLRRRPPRRSRRRPVRVRVRRRR 47
ID      Q8V7G0 PRELIMINARY; PRT; 744 AA.
AC      Q8V7G0;
DT      01-MAR-2002 (TREMBLrel. 20, Created)
DT      01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      ORF1.
OS      TT virus.
OC      Viruses; ssDNA viruses; unclassified ssDNA viruses.
CX      NCBI_TaxID=68887;
FN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CT44F;

```

```
RA Okamoto H.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CT4F;  
RX MEDLINE=21844401; PubMed=11855633;  
RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,  
PA Okamoto H.;  
RT "Analysis of the complete genomes of thirteen TT virus variants  
KT classifiable into the fourth and fifth genetic groups, isolated from  
RT viremic infants";  
RL Arch. Virol. 147:21-41(2002).  
DR EMBL: AB064605; BAB79350.1; -  
DR InterPro: IPR004219; TVirus_Unk.  
DR Pfam: PF02956; TT_ORF1; 1.  
SQ SEQUENCE 744 AA; 87602 MW; 33CD5349CFEBA69 CRC64;
```

Query Match 32.2%; Score 80.5; DB 12; Length 744;
Best Local Similarity 45.8%; Pred. No. 0.13; Indels 15; Gaps 3
Matches 22; Conservative 2; Mismatches 9;

Cy 9 WVRVRRVVRRVVRRVWR-VRVRVW-----RVRVVRVRWR 46
Db 5 WWKRRRGQWR-----RRWTGRLRRRRPQRSRRRRVRVRRRRR 47

RESULT 8
O9DUC7 PRELIMINARY; PRT; 732 AA.

ID O9DUC7 AC
OC O9DUC7;
DT 01-MAR-2001 (TReMbIrel. 16, Created)
DT 01-MAR-2001 (TReMbIrel. 16, Last sequence update)
DE 01-JUN-2002 (TReMbIrel. 21, last annotation update)
OS ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
CX NCBI_TaxID=68887;
CN [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=MF-TTV3;
RA Okamoto H.;
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MF-TTV3;
RX MEDLINE=20534983; PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
RT phylogenetic relatedness";
RL Virology 277:368-378(2000).
DL EMBL: AB041958; BAB19310.1; -
DR InterPro: IPR004219; TVirus_Unk.
DR Pfam: PF02956; TT_ORF1; 1.
SQ SEQUENCE 732 AA; 86905 MW; C28AD10850BEEA2 CRC64;

Query Match 32.0%; Score 80; DB 12; Length 732;
Best Local Similarity 43.3%; Pred. No. 0.21; Indels 26; Gaps 4
Matches 26; Conservative 1; Mismatches 7;

Cy 7 RRVVRRVVRVW-----RRVVRVVRVVRVVRVVRVVRVVRV 47
Db 6 RRWR-R-RRRRRWGRRRVGSRGVAGRRRRRRVRR--RRRRRVWR---GRAVRL 58

RESULT 9
O9YKL1 PRELIMINARY; PRT; 770 AA.

ID O9YKL1 AC
OC O9YKL1;
DT 01-MAY-1999 (TReMbIrel. 10, Created)
DT 01-MAY-1999 (TReMbIrel. 10, Last sequence update)
DE 01-JUN-2002 (TReMbIrel. 21, last annotation update)
OS ORF1 protein.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:43:01 ; Search time 44.0851 Seconds
(without alignments)
224.345 Million cell updates/sec

Title: US-10-079-075-12

Perfect score: 250
Sequence: 1 RVRVRVRVRVRVRVRVRV.....RVRVRVRVRVRVRVRV 48

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mmc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_proteob:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.5	39.4	133	5	Q9VX67
2	85	34.0	49	12	Q8V7C3
3	85	34.0	735	12	Q9DUC9
4	83	33.2	759	12	Q9IC22
5	81.5	32.6	766	12	Q9IFV0
6	80.5	32.2	49	12	Q8V7C1
7	80.5	32.2	744	12	Q8V7G0
8	80	32.0	732	12	Q9DUC7
9	80	32.0	770	12	Q9VX67
10	79	31.6	766	12	Q9VX67
11	78.5	31.4	49	12	Q8V7C2
12	78.5	31.4	171	12	Q71097
13	78.5	31.4	178	10	Q948J2
14	78.5	31.4	743	12	Q8V7G3
15	78.5	31.4	763	12	Q99A78
16	78	31.2	683	12	Q9UG64

17	78	31.2	683	12	Q9UG47	Q9UG47
18	78	31.2	720	12	Q9DUB7	Q9DUB7
19	78	31.2	767	12	Q9QUD8	Q9QUD8
20	77.5	31.0	53	12	Q8V7C0	Q8V7C0
21	77.5	31.0	53	12	Q8V7B7	Q8V7B7
22	77.5	31.0	712	12	Q9DUC1	Q9DUC1
23	77.5	31.0	726	12	Q99A80	Q99A80
24	77.5	31.0	726	12	Q8V7F7	Q8V7F7
25	77.5	31.0	746	12	Q9WAX2	Q9WAX2
26	77.5	31.0	768	12	Q70810	Q70810
27	77.5	31.0	769	12	Q70798	Q70798
28	77.5	31.0	769	12	Q70802	Q70802
29	77.5	31.0	769	12	Q70804	Q70804
30	77.5	31.0	769	12	Q9WQH0	Q9WQH0
31	77	30.8	764	12	Q9UGT1	Q9UGT1
32	77	30.8	770	12	Q70739	Q70739
33	77	30.8	770	12	Q9QUB8	Q9QUB8
34	77	30.8	770	12	Q9DUC0	Q9DUC0
35	77	30.8	770	12	Q9DHA8	Q9DHA8
36	77	30.8	770	12	Q70796	Q70796
37	77	30.8	770	12	Q70800	Q70800
38	77	30.8	770	12	Q9QUB9	Q9QUB9
39	77	30.8	770	12	Q9WQZ0	Q9WQZ0
40	77	30.8	770	12	Q9QUB3	Q9QUB3
41	76	30.4	111	10	Q39682	Q39682
42	74.5	29.8	69	12	Q8V7F2	Q8V7F2
43	74.5	29.8	739	12	Q8V717	Q8V717
44	74.5	29.8	765	12	Q9DUB8	Q9DUB8
45	74	29.6	67	12	Q8V7C6	Q8V7C6

ALIGNMENTS

RESULT 1

Q9VX67 PRELIMINARY; PRT; 133 AA.

AC Q9VX67;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C65172 protein.
GN C65172.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.B., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gilbert W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,


```
CC CC -1- SUBCELLULAR LOCATION: NUCLEAR (POSSIBLE).
CC CC -1- TISSUE SPECIFICITY: TRANSCRIBED IN ALL CELLS DURING EMBRYONIC DEVELOPMENT.
CC CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED THROUGHOUT EMBRYONIC, LARVAL, PUPAL AND ADULT STAGES AT RELATIVELY CONSTANT LEVELS.
CC CC -1- SIMILARITY: CONTAINS 13 HAT REPEATS.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
-----
DR DR EMBL; X58374; CAA11263.1; -.
DR DR EMBL; AE003423; AAF45760.1; -.
DR DR EMBL; AL009195; CAA15705.1; -.
DR DR PIR; A39634; A39634.
DR DR FlyBase; FBgn0000377; crn.
DR DR InterPro; IPR001407; HAT.
DR DR InterPro; IPR001440; TPR.
DR DR Pfam; PF02184; HAT; 10.
DR DR SMART; SMO0386; HAT; 14.
KM KM Nuclear protein; Repeat; Neurogenesis.
FT FT REPEAT 56 88 HAT 1.
FT FT REPEAT 90 122 HAT 2.
FT FT REPEAT 124 156 HAT 3.
FT FT REPEAT 158 189 HAT 4.
FT FT REPEAT 191 222 HAT 5.
FT FT REPEAT 224 259 HAT 6.
FT FT REPEAT 261 295 HAT 7.
FT FT REPEAT 305 337 HAT 8.
FT FT REPEAT 339 373 HAT 9.
FT FT REPEAT 383 419 HAT 10.
FT FT REPEAT 454 486 HAT 11.
FT FT REPEAT 488 522 HAT 12.
FT FT REPEAT 524 555 HAT 13.
FT FT DOMAIN 620 628 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT FT CONFUCT 94 94 R -> P (IN REF. 1).
SQ SQ SEQUENCE 702 AA; 84261 MW; 8E03A6B984EA412 CRC64;

Query Match          22.4%; Score 56; DB 1; Length 702;
Best Local Similarity 23.5%; Pred. No. 7.1;
Matches 16; Conservative 7; Mismatches 17; Indels 28; Gaps 3;
```

Search completed: June 9, 2003, 11:56:43
Job time : 10.7021 secs

```
CC - FUNCTION: PROTAMINIS SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
CC Sperm during the haploid phase of spermatogenesis. They compact  
CC sperm DNA into a highly condensed, stable and inactive complex.  
CC - SUBCELLULAR LOCATION: Nuclear.  
CC - TISSUE SPECIFICITY: Testis.  
CC - SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.  
DR PIR; S02786; S02786.  
DR PIR; S10754; S10754.  
KW Chromosomal protein, Nucleosome core; Spermatogenesis; DNA-binding;  
KW testis; DNA condensation; Nuclear protein.  
SQ SEQUENCE 62 AA; 8418 MW; 15EABCI09F2AD6B1 CRC64;  
  
Qy Query Match 22.4%; Score 56; DB 1; Length 62;  
Best Local Similarity 52.6%; Pred. No. 0.53; Indels 2; Gaps 1;  
Matches 20; Conservative 2; Mismatches 14;  
  
Db 11 RRVRRVVRVVRRVVRVVRV--RVRVRVVVRVVRVR 46  
|||::|||::|||::|||::|||::|||::|||  
19 RLVLRLRRRRRSSRRRRRCRRRRRRHVRVRVRRR 56  
  
RESULT 14  
YOO6 CAEEL STANDARD; PRT; 477 AA.  
AC Q09256;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 55.9 kDa protein EEBD8.6 in chromosome II.  
GN EEBD8.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol NZ;  
RA Chissey S.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC -----  
CC EMBL; U23484; AAC46766.1; -.  
DR WormPep; EEBD8.6; CE01890.  
KW Hypothetical protein.  
SQ SEQUENCE 477 AA; 55896 MW; BD671739BAF8IC58 CRC64;  
  
Qy Query Match 22.4%; Score 56; DB 1; Length 47;  
Best Local Similarity 29.4%; Pred. No. 4.7;  
Matches 10; Conservative 7; Mismatches 17; Indels 0; Gaps 0;  
  
Db 14 RRVRRVVRVVRRVVRVVRVVRVVRV 47  
|||::|||::|||::|||::|||::|||:  
366 RICWDIPFKITRSSLNSTRMWRLLVAEEKWEI 399  
  
RESULT 15  
CRN DROME STANDARD; PRT; 702 AA.  
AC P17886; Q24283; O46071;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Crooked neck protein.  
DS CRN OR EG:308b.1 OR CG3193.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
```

OC Insecta: Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=1271;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91257574; PubMed=2044955;
RA Zhang K., Smouse D., Perrimon N.;
RT "The crooked neck gene of *Drosophila* contains a motif found in a
RL family of yeast cell cycle genes.";
RN Genes Dev. 5:1080-1091(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Ament J.E., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Barton R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.C.,
RA Adair J.F., Agayanni A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernstein B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Borchen M.A., Bork J., Brokstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Paolis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Flier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostal D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalish F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Moadarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheller P., Shen H.,
RA Shne B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*";
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=20196011; PubMed=10731137;
RA Benos P.V., Gatt M.X., Ashburner M., Murphy L., Harris D.,
RA Barrett B.G., Ferriz C., Vidal S., Brunn C., Demillies J., Borokova D.,
RA Drano S., Gloux S., Leleure V., Mortier S., Galibert F., Borokova D.,
RA Mitana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA Papagannakis G., Spanos L., Cox S., Madeno E., de Pablo B.,
RA Mocololi J., Peter A., Schoettler P., Werner M., Mouritsen F.,
RA Benett N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlatsosou A., Henderson N.S.,
RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
RA Glover D.M.;
RT "From sequence to chromosome: the tip of the X chromosome of *D. melanogaster*";
RL Science 287:2220-2222(2000).
CC -I- FUNCTION: INVOLVED IN NEUROGENESIS. LOSS OF ZYGOTIC EXPRESSION
OF CN CAUSES DEFECTS IN THE PROLIFERATION OF BRAIN NEUROBLASTS
AND RESULTS IN THE ABSENCE OF IDENTIFIED NEURONAL LINEAGES IN THE
CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.

Db 750 VGEWIEEMKEEMRDVYLAKWAKATK 777

RESULT 11

RFE_PASMU ID RFE_PASMU STANDARD; PRT; 357 AA.

AC O9CNG8; 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase
DB (EC 2.4.1.-)
GN RFE OR PM0463.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxId=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=2114586; PubMed=11248100;
RA May B.J., Zhang O., Li L.L., Pautslian M.L., Whitlam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + undecaprenyl
monophosphate =UMP + undecaprenyl-N-acetyl-alpha-D-glucosaminyl
pyrophosphate.
CC -1- COFACTOR: Magnesium and manganese (By similarity).
CC -1- PATHWAY: Lipopolysaccharide biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(By similarity).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. WECA
SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AE006082; AKK02547.1; -
DR InterPro; IPR000715; Glycos transf_4.
DR Pfam; PF00953; Glycos transf_4; 1.
KM Lipopolysaccharide biosynthesis; Glycosyltransferase; Transferase;
KM Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 2 24 POTENTIAL.
FT TRANSMEM 39 60 POTENTIAL.
FT TRANSMEM 67 84 POTENTIAL.
FT TRANSMEM 126 148 POTENTIAL.
FT TRANSMEM 155 172 POTENTIAL.
FT TRANSMEM 182 204 POTENTIAL.
FT TRANSMEM 209 231 POTENTIAL.
FT TRANSMEM 241 260 POTENTIAL.
FT TRANSMEM 290 312 POTENTIAL.
FT TRANSMEM 317 339 POTENTIAL.
SQ SEQUENCE 357 AA; 40285 MW; B82A186EA21EBD5C CRC64;

Query Match 22.6%; Score 56.5; DB 1; Length 357;
Best Local Similarity 38.5%; Pred. No. 3;
Matches 10; Conservative 6; Mismatches 5; Indels 5; Gaps 1;

OY 13 VRRVRRVVRVVRVVRVVRVVRV 38
DB 333 ITRAMR-----ITRWIRMRRAKRI 353

RESULT 12
HSP3_HORSE STANDARD; PRT; 58 AA.
AC P15343;

DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm histone P2B (ST2B).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxId=9796;
RN [1]
RP SEQUENCE.
RX MEDLINE=90304188; PubMed=2364093;
RA Pirhonen A., Valtonen P., Linnala-Kankkunen A., Heiskanen M.-L.,
RA Maenpaa P.K.;
RT "Primary structures of two protamine 2 variants (St2a and St2b) from
stallion spermatozoa.";
RL Biochim. Biophys. Acta 1039:177-180(1990).
RN [2]
RP SEQUENCE OF 1-25.
RX MEDLINE=89171259; PubMed=2924903;
RA Pirhonen A., Linnala-Kankkunen A., Maenpaa P.K.;
RT "Comparison of partial amino acid sequences of two protamine 2
variants from stallion sperm. Structural evidence that the variants
are products of different genes.";
RL FEBS Lett. 244:199-202(1989).

CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COME
CC SPMEN DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.
DR PIR; S02787; S02787.
DR PIR; S10755; S10755.
KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
SQ SEQUENCE 58 AA; 7979 MW; 8D31F36098A73179 CRC64;

Query Match 22.4%; Score 56; DB 1; Length 58;
Best Local Similarity 52.6%; Pred. No. 0.5;
Matches 20; Conservative 2; Mismatches 14; Indels 2; Gaps 1;

OY 11 RRVRRVRRVVRVVRVVRVVRV 46
DB 15 RRVRRVRRVVRVVRVVRVVRV 52

RESULT 13
HSP2_HORSE STANDARD; PRT; 62 AA.
AC P15342;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm histone P2A (ST2A).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxId=9796;
RN [1]
RP SEQUENCE.
RX MEDLINE=90304188; PubMed=2364093;
RA Pirhonen A., Valtonen P., Linnala-Kankkunen A., Heiskanen M.-L.,
RA Maenpaa P.K.;
RT "Primary structures of two protamine 2 variants (St2a and St2b) from
stallion spermatozoa.";
RL Biochim. Biophys. Acta 1039:177-180(1990).
RN [2]
RP SEQUENCE OF 1-25.
RX MEDLINE=89171259; PubMed=2924903;
RA Pirhonen A., Linnala-Kankkunen A., Maenpaa P.K.;
RT "Comparison of partial amino acid sequences of two protamine 2
variants from stallion sperm. Structural evidence that the variants
are products of different genes.";
RL FEBS Lett. 244:199-202(1989).

RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=90355832; PubMed=2201868;
 RA Iuchi S., Matsuda Z., Fujiwara T.,
 RT "the arcB gene of *Escherichia coli* encodes a sensor-regulator protein
 RT for anaerobic repression of the arc modulation.",
 RL Mol. Microbiol. 4:715-727(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shaoy Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.",
 RL Science 277:1453-1474(1997).
 RN [3]
 RP CHARACTERIZATION.
 RC STRAIN=M15;
 RX MEDLINE=97431492; PubMed=9286997;
 RA Georgellis D., Lynch A.S., Lin E.C.C.;
 RT "in vitro phosphorylation study of the arc two-component signal
 RT transduction system of *Escherichia coli*.",
 RL J. Bacteriol. 179:5429-5435(1997).
 RN [4]
 RP CHARACTERIZATION.
 RC STRAIN=M15;
 RX MEDLINE=99047671; PubMed=9830034;
 RA Georgellis D., Kwon O., De Wulf P., Lin E.C.C.;
 RT "Signal decay through a reverse phosphorylay in the arc two-component
 RT signal transduction system.",
 RL J. Biol. Chem. 273:32864-32869(1998).
 RN [5]
 RP MUTAGENESIS OF HIS-292; ASP-576 AND HIS-717.
 RC STRAIN=K12 / WC4100;
 RX MEDLINE=20309722; PubMed=10851007;
 RA Kwon O., Georgellis D., Lin E.C.C.;
 RT "Phosphorelay as the sole physiological route of signal transmissison
 RT by the arc two-component system of *Escherichia coli*.",
 RL J. Bacteriol. 182:3858-3862(2000).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.06 ANGSTROMS) OF 660-778.
 RX MEDLINE=97207018; PubMed=9054511;
 RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
 RT "Insights into multistep phosphorylay from the crystal structure of
 RT the C-terminal HPC domain of ArcB.",
 RL Cell 88:717-723(1997).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 659-776 IN COMPLEX WITH CHEY
 RX MEDLINE=98437504; PubMed=9761838;
 RA Kato M., Mizuno T., Hakoshima T.;
 RT "Crystallization of a complex between a novel C-terminal transmitter,
 RT HPC domain, of the anaerobic sensor kinase ArcB and the Chemotaxis
 RT response regulator CheY.",
 RL Acta Crystallogr. D 54:140-142(1998).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 659-776.
 RX MEDLINE=20003135; PubMed=10531481;
 RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
 RT "Refined structure of the histidine-containing-phosphotransfer (HPC)
 RT domain of the anaerobic sensor kinase ArcB from *Escherichia coli* at
 RT 1.57-A resolution.",
 RL Acta Crystallogr. D 55:1842-1849(1999).
 RN [9]
 RP FUNCTION: Member of the two-component regulatory system arcB/arcA
 CC Sensor-regulator protein for anaerobic repression of the arc
 CC modulon. Activates arcA via a four-step phosphorylay. ArcB can
 CC also dephosphorylate arcB by a reverse phosphorylay involving His
 CC 717 and Asp-576.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Probable).
 CC -1- PTM: Activation requires a sequential transfer of a phosphate

[illegible]

```
CC   entities requires a license agreement (See http://www.isb-sdb.ch/announce/  
or send an email to license@isb-sdb.ch).  
-----  
CC   DR EMBL; AP001511; BABO4994.1; -.  
CC   DR InterPro; IPR001324; PRK.  
CC   DR InterPro; IPR00764; Uridine_kin.  
CC   DR Pfam; PF00485; PRK; 1.  
CC   DR PRINTS; PR00478; PHRIBLKINASE.  
CC   DR PRINTS; PFO0988; URIDINKINASE.  
CC   DR TIGRFAMs; TIGR00235; udk; 1.  
CC   KW Transferase; Kinase; ATP-binding; Complete proteome.  
FT   NP BIND           12      19  
SQ     SEQUENCE       211 AA; 24387 MW; CZAFAACB0030520B CRC64;  
  
Query Match          23.6%; Score 59; DB 1; Length 211;  
Best Local Similarity 25.6%; Pred. No. 0.84;  
Matches    11; Conservative    14; Mismatches    18; Indels    0; Gaps    0;  
  
Qy      3 VRVVRWVRWRVRVRVRVRVRVRVRVRVRVRVRVRWRM 45  
Db      137 IRIRRWVDIREGRGTLESVEIQYRVPMEMGFIEPTKY 179  
  
RESULT 9  
ARC_B_ECO57  
ID   ARC_B_ECO57        STANDARD;             PRT;              778 AA.  
AC   PS8363;  
DT   15-JUN-2002 (Rel. 41, Created)  
DT   15-JUN-2002 (Rel. 41, Last sequence update)  
DT   15-JUN-2002 (Rel. 41, Last annotation update)  
DE   Aerobic respiration control sensor protein arcB (EC 2.7.3.-).  
GN   ARC_B OR Z4574 OR ECS4089.  
OS   Escherichia coli O157:H7.  
OC   Bacteria; Proteobacteriae; gamma subdivision; Enterobacteriaceae;  
OX   Escherichia.  
NCBI_TaxID=83334;  
[1]  
RP     SEQUENCE FROM N.A.  
RC     STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX     MEDLINE=21074935; PubMed=11206551;  
RA     Pena N.T., Plunkett G. II, Burland V., Mau B., Glaesner J.D.,  
RA     Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA     Grofbeck E.J., Davis N.W., Lim A., Dimatanta E.T., Potamousis K.,  
RA     Apodaca J., Anantharaman T.S., Lih Y., Yen G., Schwartz D.C.,  
RA     Welch R.A., Blattner F.R.;  
RL     "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7,"  
RT     Nature 409:552-553(2001).  
RN     [2]  
RP     SEQUENCE FROM N.A.  
RC     STRAIN=O157:H7 / RIMD 0509952;  
RX     MEDLINE=21156231; PubMed=11258796;  
RA     Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ichii K., Yokoyama K.,  
RA     Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,  
RA     Tida T., Takami H., Honda T., Saekawa C., Ogasawara N., Yasunaga T.,  
RA     Kuhara S., Shibata T., Hattori M., Shimagawa H.;  
RT     "Complete genome sequence of enterohemorrhagic Escherichia coli  
OT   O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL     DNA Res. 8:11-22(2001).  
CC     -| FUNCTION: Member of the two-component regulatory system arcB/arcA.  
CC         Sensor-regulator protein for anaerobic repression of the arc  
CC         modulon. Activates arcA via a four-step phosphorylay. ArcB can  
CC         also dephosphorylate arcA by a reverse phosphorelay involving His-  
CC         717 and Asp-576 (By similarity).  
CC     -| SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC         (Probable).  
CC     -| PTM: Activation requires a sequential transfer of a phosphate  
CC         group from a His in the primary transmitter domain, to a Asp in  
CC         the receiver domain and to a His in the secondary transmitter  
CC         domain (By similarity).  
CC     -| SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.  
CC     -| SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.  
CC     -| SIMILARITY: CONTAINS 1 PAS (PBR-AENT-SIM) DIMERIZATION DOMAIN.
```

[illegible]

OS *Drosophila melanogaster* (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkelley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Borkov A., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkllov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RU Science 287:2185-2195(2000).
 RN [2]
 RP IDENTIFICATION, AND TISSUE SPECIFICITY.
 RX MEDLINE=20175760; PubMed=10710312;
 RA Clyne P.J., Warr C.G., Carlson J.R.;
 RT "Candidate taste receptors in *Drosophila*."
 RU Science 287:1830-1834(2000).
 RN [3]
 RP IDENTIFICATION.
 RX MEDLINE=21407712; PubMed=11516643;
 RA Dunipace L., Meister S., McNealy C., Amrein H.;
 RT "Spatially restricted expression of candidate taste receptors in the
 RT *Drosophila* gustatory system."
 RU Curr. Biol. 11:822-835(2001).
 RN [4]
 RP CONCEPTUAL TRANSLATION.
 RA Robertson H.;
 RL Unpublished observations (NOV-2001).
 CC -1- FUNCTION: Probable role in the gustatory response.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Expressed in the adult labellar chemosensory
 CC neurons.
 CC -1- SIMILARITY: BELONGS TO FAMILY DR-TR OF G-PROTEIN COUPLED
 CC RECEPTORS. SUBFAMILY VI.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
 CC gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AE003459; AAF46958.2; ALT_SEQ.
 CC DR FlyBase; FBgn0041235; Gr59c.
 CC KW Hypothetical protein; Receptor; G-protein coupled receptor;
 CC Transmembrane; Glycoprotein; Multigene family.
 CC FT DOMAIN 1 3
 CC FT TRANSMEM 4 24
 CC FT DOMAIN 25 39
 CC FT TRANSMEM 40 60
 CC FT DOMAIN 61 75
 CC FT TRANSMEM 76 96
 CC FT DOMAIN 97 166
 CC FT TRANSMEM 167 187
 CC FT DOMAIN 188 259
 CC FT TRANSMEM 260 280
 CC FT DOMAIN 281 284
 CC FT TRANSMEM 285 305
 CC FT DOMAIN 306 372
 CC FT TRANSMEM 373 393
 CC FT DOMAIN 394 397
 CC FT CARBOHYD 61 61
 CC SQ SEQUENCE 397 AA; 46164 MW; 60CFEC8AC46F90E CRC64;
 CC
 CC Query Match 24.4%; Score 61; DB 1; Length 397;
 CC Best Local Similarity 30.8%; Pred. No. 0.94;
 CC Matches 12; Conservative 10; Mismatches 7; Indels 10; Gaps 2;
 CC
 CC QY 2 VARVARRRRRR-RVRRRVRRVRR-----RVRR 30
 CC DB 91 LSLITRWQSRFRIRIMQIALVDRPDQVGRGWR 129
 CC
 CC RESULT 8
 CC ID URK_BACHD STANDARD; PRT; 211 AA.
 CC AC G9XK08;
 CC DT 15-JUN-2002 (Rel. 41, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Uridine kinase (EC 2.7.1.48). (Uridine monophosphokinase) (Cytidine
 CC monophosphokinase).
 CC GN UDK OR BH1275.
 CC OS Bacillus halodurans.
 CC OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC OX NCBI_TaxID=86665;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=C-125 / JCM 9153;
 CC RX MEDLINE=20512582; PubMed=11058132;
 CC RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 CC Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 CC Horikoshi K.;
 CC RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 CC RT *halodurans* and genomic sequence comparison with *Bacillus subtilis*."
 CC RL Nucleic Acids Res. 28:4317-4331(2000).
 CC CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

```
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on li-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC
DR DR EMBL; U70921; AAC83411.1; -.
DR InterPro; IPR004912; Adeno_VII.
DR Pfam; PF03228; Adeno_VII; 1.
FT FT PROPEP 1 24 BY SIMILARITY.
FT CHAIN 25 193 MAJOR CORE PROTEIN.
FT SITE 24 25 CLEANAGE (BY ADENOVIRUS PROTEASE)
FT SEQUENCE 193 AA; 21358 MW; 43137E07DB379DD0 CRC64;
Cc Cc
Query Match 25.0%; Score 62.5; DB 1; Length 193;
Best Local Similarity 40.4%; Pred. No. 0.29;
Matches 21; Conservative 6; Mismatches 16; Indels 9; Gaps 3;
Cc Cc
QY 2 VVRVVRWVR--RVRRVWR-----VVVRVWRVRRVRRVWR-VVRVVR 44
Db 88 VVSDDRARRAPAKSRRRRIARRRHSRTMARAAAPALRRARRCRGRAMLAARR 139
Cc Cc
RESULT 5
ERMS_STRFR
ID ERMS_STRFR STANDARD; PRT; 319 AA.
AC P45439;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA adenine N-6-methyltransferase (EC 2.1.1.48) (Macrolide-
DE lincosamide-streptogramin B resistance protein) (Erythromycin
DE resistance protein).
GN ERMSF OR TURK.
OS Streptomyces fradiae.
OC Bacteria; Actinobacteria; Actinobacteriales (class); Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxId=1906;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RC MEDLINE=88169508; PubMed=3127381;
RA Kamimura S., Weisblum B.;
RT "Translational attenuation control of ermsF, an inducible resistance
RT determinant encoding RNA N-methyltransferase from Streptomyces
RT fradiae.";
RL J. Bacteriol. 170:1800-1811(1988).
CC CC -1- FUNCTION: THIS PROTEIN PRODUCES A DIMETHYLATION OF THE ADENINE
CC RESIDUE AT POSITION 2058 IN 23S rRNA, RESULTING IN REDUCED
CC AFFINITY BETWEEN RIBOSOMES & MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B
CC ANTIBIOTICS.
CC CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing N(6)-methyladenine.
CC CC -1- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
CC FAMILY.
Cc Cc
-----
Cc Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on li-
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to licenses@isb-sib.ch).
Cc
DR DR EMBL; M19269; AAA26742.1; -.
DR InterPro; IPR001737; RNA_A_dimeth.
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF00398; RtmAD; 1.
DR PROSITE; PS01131; RNA_A_DIMETH; 1
```

```

SQ      Antibiotic resistance; Transferrase; Methyltransferase.
SEQUENCE   319 AA; 35527 MW; 3A543FA222CFB7DB CRC64;

Query Match          25.0%; Score 62.5; DB 1; Length 319;
Best Local Similarity 43.9%; Pred. No. 0.49;
Matches 18; Conservative 4; Mismatches 16; Indels 3; Gaps 1;

Oy      VRRVRRRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVR 46
Db      203 VRW---PRHEMRLGRVSRREPRFVRVDSGILIRRRER 240

RESULT 6
VC07_ADE40
ID_VC07_ADE40 STANDARD; PRT; 165 AA.
AC Q89532;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
OS Human adenovirus type 40.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxId=28284;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RX MEDLINE=94087748; PubMed=8263936;
RA Davison A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
RT "The DNA sequence of adenovirus type 40."
RN J. Mol. Biol. 234:1308-1316(1993).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RA Pieniazek N.J., Slemenda S.B., Pieniazek D., Luftig R.B.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).

CC CC
CC EMBL; L19443; AAC13963.1; -.
CC DR EMBL; M8665; AAA42526.1; -.
CC DR InterPro; IPR004912; Adeno_VII.
CC PR Pfam; PF03228; Adeno_VII; 1.
CC FT PROPEP 1 BY SIMILARITY.
CC FT CHAIN 24 MAJOR CORE PROTEIN.
CC FT SITE 23 CLEAVAGE (BY ADENOVIRUS PROTEASE)
CC FT SITE 24 (POTENTIAL).
SO SEQUENCE 185 AA; 20518 MW; 4PB0B53EF218ABE CRC64;

Query Match          24.4%; Score 61; DB 1; Length 185;
Best Local Similarity 35.6%; Pred. No. 0.42;
Matches 16; Conservative 8; Mismatches 15; Indels 6; Gaps 1;

Oy      2 VRRVRRRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVR 40
Db      87 VNANARYAQRKRRLQRRRRRPFATMAARAVALFRRAQRIGRRMR 131

RESULT 7
G59C_DROME
ID_G59C_DROME STANDARD; PRT; 397 AA.
AC Q9W1U5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative gustatory receptor 59c.
NC GR59C OR GR59D.2 OR CG13543 OR CG30186.
```


	ID	_YACG_RHISN	STANDARD;	PRT;	305 AA.
AC	P55389;				
DT	01-NOV-1997	(Rel. 35,			
DJ	01-NOV-1997	(Rel. 35,			
DR	15-JUN-2002	(Rel. 41,			
DE	Probable DNA-invertase Y4CG.				
GN	Y4CG.				
OS	Rhizobium sp. (strain NGR234).				
OC	Plasmid bym pNGR234e.				
CC	Bacteria; Proteobacteriae; alpha subdivision; Rhizobiaceae group;				
OC	Rhizobiaceae; Rhizobium.				
OX	NCBI_TaxID=394;				
RP	[1]				
FP	SEQUENCE FROM N.A.				
RX	MEDLINE=97305956; PubMed=9163424;				
RA	Freiburg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,				
RL	Perrer X ;				
CC	"Molecular basis of symbiosis between Rhizobium and legumes."				
CC	Nature 387:394-401(1997)				
CC	-1- SIMILARITY: BELONGS TO THE SITE-SPECIFIC RECOMBINASE RESOLVAS				
CC	FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/a				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AE000068; AAB92422.1;				
DR	HSSP; P03012; 2RSU.				
DR	InterPro; IPR001822; Recombinase.				
DR	Pfam; PF00239; resolvasc; 1.				
DR	PROSITE; PS00397; RECOMBINASES_1; 1.				
DR	PROSITE; PS00398; RECOMBINASES_2; 1.				
KW	Hypothetical protein; DNA recombination; DNA integration; DNA-binding;				
FT	DNA invertase; Plasmid.				
FM	ACT_SITE 23 .. 23				
FT	TRANSIENT COVALENT LINKAGE TO DNA DURING				
FT	STRAND CLEAVAGE AND REJOINING				
FT	(BY SIMILARITY).				
SQ	SEQUENCE 305 AA; 34277 MM; 48C03BD3A4A9420F CRC64;				
Query Match		25.2%;	Score 63;	DB 1;	Length 305;
Best Local Similarity		40.5%;	Pred. No. 0.4;		
Matches 17; Conservative		6;	Mismatches	9;	Indels 10;
Gaps					3;
Oy	9 WVRRVRVVRR-----WVRRVVRRVVRRVVRR 40	:: :	:: :	:: :	:: :
Db	186 WLPTVRLRLPRHSMDNVVRIILNRGHDMVTVERLRPAVRLRV 227	:: :	: :	:	:
RESULT 3					
y278_MYCTU					
ID y278 MYCTU	STANDARD;	PRT;	957 AA.		
AC P56877;					
DT 30-MAY-2000	(Rel. 39,				
DT 30-MAY-2000	(Rel. 39,				
DT 16-OCT-2001	(Rel. 40,				
DE Hypochemical PE-PGRS family protein RV0278c precursor.					
GN RV0278C OR MT0291 OR MTV035.06C.					
OS Mycobacterium tuberculosis.					
OC Bacateria; Actinobacteria; Actinobacteriae (class); Actinobacteridae;					
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.					
OX NCBI_TaxID=1773;					
RM [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN-H37RV;					
RX MEDLINE=98293987; PubMed=9634230;					
RA Cole S.T.; Brosch R.; Parkhill J.; Garnier T.; Churcher C.; Harris D.,					

[illegible]

US-08-932-682-160

Query Match 26.3%; Score 50; DB 2; Length 28;
Best Local Similarity 42.3%; Pred. No. 5.9;
Matches 11; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 7 RVRVVRVRVRVRVRVRVRVRVR 32
Db 1 RVIRVVOGACRAIRHPRIRIOGLRR 26

Search completed: June 9, 2003, 12:05:10
Job time : 13.2553 secs

MOLECULE TYPE: protein
US-08-980-357-28

Query Match 26.8%; Score 51; DB 3; Length 2254;
Best Local Similarity 45.5%; Pred. No. 2.5e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 14 RRVRRVRRVRRVRRVRRVRR 35
DB 1627 RAVRSVVRVYVETVPRLLARWTK 1648

RESULT 13
US-09-199-637A-289
; Sequence 289, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 289
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-289

Query Match 26.6%; Score 50.5; DB 4; Length 101;
Best Local Similarity 43.2%; Pred. No. 17;
Matches 16; Conservative 0; Mismatches 16; Indels 5; Gaps 2;

QY 2 RRVRRVRRVRRVRRVRRVRR 33
DB 44 RRSRRSSRPSRRRRRGARATSHAPASSTRPSTW 80

RESULT 14
US-08-786-748A-160
; Sequence 160, Application US/08786748A
; Patent No. 5714577
; GENERAL INFORMATION:
; APPLICANT: Ronald, Montelaro C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Mieczner, Timothy A.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,748A

FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-160

Query Match 26.3%; Score 50; DB 1; Length 28;
Best Local Similarity 42.3%; Pred. No. 5.9;
Matches 11; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 7 RRVRRVRRVRRVRRVRRVRRVRR 32
DB 1 RVRVVGACRAIRHPRIRIROGLRR 26

RESULT 15
US-08-932-682-160
; Sequence 160, Application US/08932682
; Patent No. 5945507
; GENERAL INFORMATION:
; APPLICANT: Ronald, Montelaro C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Mieczner, Timothy A.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTS, L.L.P.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,682
; FILING DATE: 18-SEP-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/786,748
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Rochelle K. Seide
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: AP30421-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-705-5000
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5945507e

RESULT 5
US-08-839-624-27
Sequence 27, Application US/08839624
Patent No. 6225045
GENERAL INFORMATION:
APPLICANT: Karm et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATTING
TITLE OF INVENTION: HIV INFECTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,624
FILING DATE: April 15, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/78191
FILING DATE: 15-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-839-624-27
Query Match 28.2%; Score 53.5; DB 4; Length 31;
Best Local Similarity 44.8%; Pred. No. 2.6;
Matches 13; Conservative 0; Mismatches 11; Indels 5; Gaps 1;
QY 2 RRVRRVVRVRRV-----RRVRRVRR 25
DB 2 RRAARRARRARRCGVSARRARRARRR 30
RESULT 6
US-09-150-812-27
Sequence 27, Application US/09150812
Patent No. 6395891
GENERAL INFORMATION:
APPLICANT: Karm et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATTING
TITLE OF INVENTION: HIV INFECTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-150-812-27
Query Match 28.2%; Score 53.5; DB 4; Length 31;
Best Local Similarity 44.8%; Pred. No. 2.6;
Matches 13; Conservative 0; Mismatches 11; Indels 5; Gaps 1;
QY 2 RRVRRVVRVRRV-----RRVRRVRR 25
DB 2 RRAARRARRARRCGVSARRARRARRR 30
RESULT 7
US-09-413-814-78
Sequence 78, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bioecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
TITLE OF INVENTION: heteropolypeptide compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 78
LENGTH: 882
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-78
Query Match 27.6%; Score 52.5; DB 4; Length 882;
Best Local Similarity 53.1%; Pred. No. 71;
Matches 17; Conservative 2; Mismatches 12; Indels 1; Gaps 1;

ADDRESSEE: Benita J. Rohm, Esq.
STREET: 6601 Woodward Avenue
CITY: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44mb, 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6;
SOFTWARE: ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436, 703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7MK-060548-00233
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-17

Query Match 30.5%; Score 58; DB 2; Length 38;
Best Local Similarity 51.9%; Pred. No. 1;
Matches 14; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 6 RRVRRVRRVRRVRRVRRVRRVRR 32
DB 3 RRAARRARRARRARRARRARRARR 29

RESULT 3
US-08-436-703B-5
Sequence 5, Application US/08436703B
Patent No. 5919761
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
APPLICANT: Stanley, James C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR
TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
TITLE OF INVENTION: WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 6601 Woodward Avenue
CITY: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44mb, 3.5"

COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6;
SOFTWARE: ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436, 703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7MK-060548-00233
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-5

Query Match 30.5%; Score 58; DB 2; Length 39;
Best Local Similarity 51.9%; Pred. No. 1;
Matches 14; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 6 RRVRRVRRVRRVRRVRRVRRVRR 32
DB 3 RRAARRARRARRARRARRARRARR 29

RESULT 4
US-08-995-172-1
Sequence 1, Application US/08995172B
Patent No. 6218112
GENERAL INFORMATION:
APPLICANT: Thatcher, David R
APPLICANT: Wilks, Paula E
TITLE OF INVENTION: Optimization of Gene Delivery and Gene Delivery Systems
FILE REFERENCE: CAC00026
CURRENT APPLICATION NUMBER: US/08/995, 172B
CURRENT FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/033, 908
EARLIER FILING DATE: 1996-12-23
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 31
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Peptide
US-08-995-172-1

Query Match 28.2%; Score 53.5; DB 4; Length 31;
Best Local Similarity 44.8%; Pred. No. 2, 6; 11; Indels 5; Gaps 1;
Matches 13; Conservative 0; Mismatches

QY 2 RRVRRVRRVRRVRRVRRVRRVRR 25
DB 2 RRAARRARRARRARRARRARRARR 30


```

1  APPLICANT: Xu, H.
2  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
3  FILE REFERENCE: EITTRA 034A
4  CURRENT APPLICATION NUMBER: US/10/282,122A
5  CURRENT FILING DATE: 2003-02-20
6  PRIOR APPLICATION NUMBER: 60/1191,078
7  PRIOR FILING DATE: 2000-03-21
8  PRIOR APPLICATION NUMBER: 60/206,848
9  PRIOR FILING DATE: 2000-05-23
10 PRIOR APPLICATION NUMBER: 60/207,727
11 PRIOR FILING DATE: 2000-05-26
12 PRIOR APPLICATION NUMBER: 60/230,335
13 PRIOR FILING DATE: 2000-09-06
14 PRIOR APPLICATION NUMBER: 60/230,347
15 PRIOR FILING DATE: 2000-09-09
16 PRIOR APPLICATION NUMBER: 60/242,578
17 PRIOR FILING DATE: 2000-10-23
18 PRIOR APPLICATION NUMBER: 60/251,625
19 PRIOR FILING DATE: 2000-11-27
20 PRIOR APPLICATION NUMBER: 60/257,931
21 PRIOR FILING DATE: 2000-12-22
22 PRIOR APPLICATION NUMBER: 60/267,636
23 PRIOR FILING DATE: 2001-02-09
24 PRIOR APPLICATION NUMBER: 60/265,308
25 PRIOR FILING DATE: 2001-02-16
26 Remaining Prior Application data removed - See File Wrapper or PALM.
27 NUMBER OF SEQ ID NOS: 78614
28 SOFTWARE: PatentIn version 3.1
29 SEQ ID NO 49117
30 LENGTH: 360
31 TYPE: PRT
32 ORGANISM: Burkholderia fungorum
33 FEATURE:
34 NAME/KEY: MISC FEATURE
35 LOCATION: (270)..(270)
36 OTHER INFORMATION: X-any amino acid
37 FEATURE:
38 NAME/KEY: MISC FEATURE
39 LOCATION: (272)..(272)
40 OTHER INFORMATION: X-any amino acid
41 FEATURE:
42 NAME/KEY: MISC FEATURE
43 LOCATION: (342)..(342)
44 OTHER INFORMATION: X-any amino acid
45 US-10-282-122A-49117

```

Query Match	32.6%;	Score 62;	DB 6;	Length 360;
Best Local Similarity	36.4%;	Pred. No. 1.2e+02;		
Matches 16; Conservative	4;	Mismatches 12;	Indels 12;	Gaps 2;

Oy 2 RRYVRRVVVRVRRV-----RVW---RVVRRVVRW 33
 ||::| |::||
Db 310 RLWQSRCPQRRLRPDGRGCGCRPAWRPXSRRQPRAVGW 353

Search completed: June 9, 2003, 12:31:39
Job time : 47.5957 secs


```

; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, UTN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12598
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Streptomyces avermectilis
; US-10-156-761-12598

Query Match          33.7%; Score 64; DB 6; Length 692;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 13; Conservative 6; Mismatches 3; Indels 4; Gaps 1;

QY      8 VVRVRRVRR---VRRVRRVRRV 29
DB      48 VRLRVRRVRRPRLPVRLVRRRIQL 73

RESULT 12
US-10-425-114-53360
; Sequence 53360, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53360
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700155612_FLI.pep
; US-10-425-114-53360

Query Match          32.6%; Score 62; DB 6; Length 107;
Best Local Similarity 48.7%; Pred. No. 46;
Matches 19; Conservative 2; Mismatches 6; Indels 12; Gaps 3;

QY      8 VVR---VVRVRRVRRVRRVRR---VRRVRR 36
DB      36 VVRBAVVGRR--RVRRRRDVRRRRRAVGLARRVRR 72

RESULT 13
US-10-425-114-56955
; Sequence 56955, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E

```

```

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56955
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3245-483-D5_FLI.pep
; US-10-425-114-56955

Query Match          32.6%; Score 62; DB 6; Length 107;
Best Local Similarity 48.7%; Pred. No. 46;
Matches 19; Conservative 2; Mismatches 6; Indels 12; Gaps 3;

QY      8 VVR---VVRVRRVRRVRRVRR---VRRVRR 36
DB      36 VVRBAVVGRR--RVRRRRDVRRRRRAVGLARRVRR 72

RESULT 14
US-10-425-114-52367
; Sequence 52367, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52367
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4073-017-F8_FLI.pep
; US-10-425-114-52367

Query Match          32.6%; Score 62; DB 6; Length 139;
Best Local Similarity 43.9%; Pred. No. 57;
Matches 18; Conservative 5; Mismatches 8; Indels 10; Gaps 3;

QY      3 RVRRVRRVRRVRRVRR---VRRVRR---VRRVRR 36
DB      7 RRRV---RRLRRVPRRLRRRLVRRRRRLRRMQPRRRRLRQ 44

RESULT 15
US-10-282-122A-49117
; Sequence 49117, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```

US-09-978-825-24220

Query Match	35.0%;	Score 66.5;	DB 5;	Length 96;
Best Local Similarity	45.2%;	Pred. No. 16;		
Matches 14;	Conservative 3;	Mismatches 13;	Indels 1;	Gaps 1;

```
QY      3 RWMRRVVRVVRVVRRVVR-RRVMRRVVRVVR 32
          ||||| | :| ||| :|||
Db     11 RCWRRVSNTRRPWAAKLCSRTWRTESTLVRR 41
```

RESULT 7
US-10-057-498-24220

```

: Sequence 24220, Application US/10057498
: GENERAL INFORMATION:
: APPLICANT: Mitcham, Jennifer
: APPLICANT: Skeiky, Yasir
: APPLICANT: Persing, David
: TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
: FILE REFERENCE: 210121.514
: CURRENT APPLICATION NUMBER: US/10/057,498
: CURRENT FILING DATE: 2001-04-20
: NUMBER OF SEQ ID NOS: 29212
: SEQ ID NO 24220
: LENGTH: 96
: TYPE: PRT
: ORGANISM: Propioni acnes
: US-10-057-498-24220

```

Query Match	35.0%;	Score 66.5;	DB 6;	Length 96;
Best Local Similarity	45.2%;	Pred. No. 16;		
Matches 14;	Conservative 3;	Mismatches 13;	Indels 1;	Gaps 1;

QY 3 RWMRRVVRVVRVVRVVRV-RVVRVVRVVRVVR 32
| | | | : | | | : | | |
Db 11 RCMRRVSNTRRPMAKLCSTWRRTESLYR 41

RESULT 8
US-10-219-051B-8227

```

Sequence 8227, Application US/10219051B
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation doing business as Massachusetts General
APPLICANT: Hospital / Bayer AG
TITLE OF INVENTION: Nucleotide sequences involved in pain
FILE REFERENCE: Lea 35693 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/219,051B
CURRENT FILING DATE: 2003-05-09
PRIORITY APPLICATION NUMBER: US 60/312,147
PRIORITY FILING DATE: 2001-08-14
PRIORITY APPLICATION NUMBER: US 60/346,382
PRIORITY FILING DATE: 2001-11-01
PRIORITY APPLICATION NUMBER: US 60/333,347
PRIORITY FILING DATE: 2001-11-26
NUMBER OF SEQ. ID NOS: 14715
SOFTWARE: Perl script
SEQ ID NO 8227
LENGTH: 396
TYPE: PRT
ORGANISM: Rattus norvegicus
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: SWISS-Prot / AAA68695
DATABASE ENTRY DATE: 1998-11-01
IS-10-219-051B-8227

```

Query Match	34.5%	Score 65.5;	DB 6;	Length 396;
Best Local Similarity	27.9%	Pred. NO. 61;		
Matches 12; Conservative	7;	Mismatches 13;	Indels 11;	Gaps 1;

```

Oy      5 WRKRVV-----VRWVRVRVRVRVRVRVRVRVRVR 36
          |:::|:::|:::|:::|:::|:::|:::|
Db      87 WKSIAKLCRCQETIANLERWVKREHMVWREVFYRLERADR 129

```

RESULT 9
US-10-219-051B-8229
; Sequence 8229, App

```

GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation doing business as Massachusetts General
APPLICANT: Hospital / Bayer AG
TITLE OF INVENTION: Nucleotide sequences involved in pain
FILE REFERENCE: Lea 35693 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/219,051B
CURRENT FILING DATE: 2003-05-09
PRIOR APPLICATION NUMBER: US 60/312,147
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: US 60/346,382
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/333,347
PRIOR FILING DATE: 2001-11-26
NUMBER OF SEQ ID NOS: 14715
SOFTWARE: Perl script
SEQ ID NO 8229
LENGTH: 396
TYPE: PRT
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Refseq / NP_056008
DATABASE ENTRY DATE: 2002-10-31
US-10-219-051B-8229

```

Query Match	34.5%	Score 65.5;	DB 6;	Length 396;
Best Local Similarity	27.9%;	Pred. No. 61;		
Matches 12; Conservative	7;	Mismatches 13;	Indels 11;	Gaps 1;

Qy 5 WRKRVRV-----VRKRVRVRVKRVRVKRVKRV 36
::: : ||| : |||
Db 87 WKSIKACLCRCQETIANLERWVKREHMYRVRYFLERWADR 129

RESULT 10
US-10-425-114-58367

```

Sequence 58367, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 58367
LENGTH: 129
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3591-091-B12_F11.pep
US-10-425-114-58367

```

Query Match	34.2%;	Score 65;	DB 6;	length 129;
Best Local Similarity	46.9%;	Pred. No. 28;		
Matches 15; Conservative	2;	Mismatches 3;	Indels 12;	Gaps 2;

[illegible]

RESULT 11
US-10-156-761-12598
; Sequence 12598, Application US/10156761

```

; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (90)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-2376
```

Query Match 38.2%; Score 72.5; DB 6; Length 136;

Best Local Similarity 38.2%; Pred. No. 5.7;

Matches 13; Conservative 11; Mismatches 9; Indels 1; Gaps 1;

```

Qy 4 RVRVRVVRVVRVVRVVRVVRVVRVVRVVR 36
    |||||:|||||:|||||:|||||:|||||
Db 47 LWRNRLRLILIMRRRLMLILRLRRMLLR 80
```

RESULT 3

US-10-366-683-32538

; Sequence 32538, Application US/10366683

; GENERAL INFORMATION:

; APPLICANT: Rubenfield, Marc J.

; APPLICANT: Noelling, Jock

; APPLICANT: Deloughery, Craig

; APPLICANT: Bush, David

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: PATH03-04

; CURRENT APPLICATION NUMBER: US/10/366,683

; CURRENT FILING DATE: 2003-02-13

; PRIOR APPLICATION NUMBER: 09/252,991

; PRIOR FILING DATE: 1999-02-18

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 32538

; LENGTH: 205

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-10-366-683-32538

Query Match 35.8%; Score 68; DB 6; Length 205;

Best Local Similarity 48.7%; Pred. No. 21;

Matches 19; Conservative 1; Mismatches 13; Indels 6; Gaps 2;

```

Qy 2 RRVRRVVRVVRVVRVVRVVRVVRVVRVVR 36
    |||||:|||||:|||||:|||||:|||||
Db 159 RRLMRARRRRRGSGVRCGWMRTTTPRGRTAGR--RRWRPR 195
```

RESULT 4

US-10-419-128-32538

; Sequence 32538, Application US/10419128

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196,136

; CURRENT APPLICATION NUMBER: US/10/419,128

; CURRENT FILING DATE: 2003-04-21

; PRIOR APPLICATION NUMBER: US/09/252,991

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 32538

; LENGTH: 205

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-10-419-128-32538

Query Match 35.8%; Score 68; DB 6; Length 205;

Best Local Similarity 48.7%; Pred. No. 21;

Matches 19; Conservative 1; Mismatches 13; Indels 6; Gaps 2;

```

Qy 2 RRVRRVVRVVRVVRVVRVVRVVRVVRVVR 36
    |||||:|||||:|||||:|||||:|||||
Db 159 RRLMRARRRRRGSGVRCGWMRTTTPRGRTAGR--RRWRPR 195
```

RESULT 5

PCT-US02-32727-24220

; Sequence 24220, Application PC/TUS0232727

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; APPLICANT: Bhatia, Ajay

; APPLICANT: Maisonneuve, Jean Francois

; APPLICANT: Zhang, Yanni

; APPLICANT: Wang, Siqing

; APPLICANT: Jen, Shyian

; APPLICANT: Lodes, Michael

; APPLICANT: Benson, Darin

; APPLICANT: Jones, Robert

; APPLICANT: Carter, Darrick

; APPLICANT: Barth, Brenda

; APPLICANT: Douglass, John

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes

; FILE REFERENCE: 210121,514C1

; CURRENT APPLICATION NUMBER: PCT/US02/32727

; CURRENT FILING DATE: 2002-10-11

; NUMBER OF SEQ ID NOS: 30992

; SEQ ID NO 24220

; LENGTH: 96

; TYPE: PRT

; ORGANISM: Propionibacterium acnes

PCT-US02-32727-24220

Query Match 35.0%; Score 66.5; DB 1; Length 96;

Best Local Similarity 45.2%; Pred. No. 16;

Matches 14; Conservative 3; Mismatches 13; Indels 1; Gaps 1;

```

Qy 3 RVRVRVVRVVRVVRVVRVVRVVRVVRVVR 32
    |||||:|||||:|||||:|||||:|||||
Db 11 RCMRRVSNTRPRWPAKLCSRTWRRTESLYVR 41
```

RESULT 6

US-09-978-825-24220

; Sequence 24220, Application US/09978825

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; APPLICANT: Bhatia, Ajay

; APPLICANT: Maisonneuve, Jean Francois

; APPLICANT: Zhang, Yanni

; APPLICANT: Wang, Siqing

; APPLICANT: Jen, Shyian

; APPLICANT: Lodes, Michael

; APPLICANT: Benson, Darin

; APPLICANT: Jones, Robert

; APPLICANT: Carter, Darrick

; APPLICANT: Barth, Brenda

; APPLICANT: Douglass, John

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes

; FILE REFERENCE: 210121,514C1

; CURRENT APPLICATION NUMBER: US/09/978,825

; CURRENT FILING DATE: 2003-01-29

; NUMBER OF SEQ ID NOS: 30992

; SEQ ID NO 24220

; LENGTH: 96

; TYPE: PRT

; ORGANISM: Propionibacterium acnes

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:57:02 ; Search time 46.5957 Seconds
(without alignments)
173.908 Million cell updates/sec

Title: US-10-079-075-11

Sequence: 1 VRRVRRVVRVVRVVRVVRVVRVVRVVRVVR 36

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1046584 seqs, 225093350 residues

Total number of hits satisfying chosen parameters: 1046584

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Parents AA New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76.5	40.3	245	US-10-425-114-70663	Sequence 70663, A
2	72.5	38.2	136	US-10-264-237-2376	Sequence 2376, Ap
3	68	35.8	205	US-10-366-683-32538	Sequence 32538, A
4	66.5	35.0	96	US-10-419-128-32538	Sequence 32538, A
5	66.5	35.0	96	PCT-US02-32727-24220	Sequence 24220, A
6	66.5	35.0	96	US-09-978-825-24220	Sequence 24220, A
7	66.5	35.0	96	US-10-057-498-24220	Sequence 24220, A
8	65.5	34.5	366	US-10-219-0518-8227	Sequence 8227, Ap
9	65.5	34.5	366	US-10-219-0518-8229	Sequence 8229, Ap
10	65	34.2	129	US-10-425-114-58367	Sequence 58367, A
11	64	33.7	692	US-10-156-761-12598	Sequence 12598, A
12	62	32.6	107	US-10-425-114-53360	Sequence 53360, A
13	62	32.6	107	US-10-425-114-56955	Sequence 56955, A
14	62	32.6	139	US-10-425-114-52367	Sequence 52367, A
15	62	32.6	360	US-10-282-122A-49117	Sequence 49117, A
16	62	32.6	360	US-10-282-122A-49117	Sequence 49117, A
17	61.5	32.4	387	US-10-366-683-16956	Sequence 16956, A
18	61.5	32.4	387	US-10-366-683-16956	Sequence 16956, A
19	61	32.1	77	PCT-US02-32727-15057	Sequence 15057, A
20	61	32.1	77	US-09-978-825-15057	Sequence 15057, A
21	61	32.1	77	US-10-057-498-15057	Sequence 15057, A
22	61	32.1	957	US-10-282-122A-64361	Sequence 64361, A
23	59.5	31.3	190	US-10-425-114-70810	Sequence 70810, A
24	59.5	31.3	195	US-10-425-114-68513	Sequence 68513, A
25	59.5	31.3	209	US-10-425-114-68513	Sequence 68513, A
26	59.5	31.3	252	US-10-425-114-71061	Sequence 71061, A

27	59.5	31.3	342	US-10-425-114-68594	Sequence 68594, A
28	58.5	30.8	152	US-10-425-114-54667	Sequence 54667, A
29	58.5	30.8	288	US-10-369-493-17910	Sequence 17910, A
30	58	30.5	66	PCT-US02-32727-12229	Sequence 12229, A
31	58	30.5	66	US-09-978-825-12229	Sequence 12229, A
32	58	30.5	66	US-10-057-498-12229	Sequence 12229, A
33	58	30.5	142	US-60-452-680-16830	Sequence 16830, A
34	57.5	30.3	66	US-10-203-138A-10862	Sequence 10862, A
35	57	30.0	53	PCT-US02-32727-20271	Sequence 20271, A
36	57	30.0	53	PCT-US02-32727-20271	Sequence 20271, A
37	57	30.0	53	PCT-US02-32727-20271	Sequence 20271, A
38	57	30.0	53	US-09-978-825-25676	Sequence 25676, A
39	57	30.0	53	US-09-978-825-25676	Sequence 25676, A
40	57	30.0	53	US-09-978-825-25676	Sequence 25676, A
41	57	30.0	53	US-10-057-498-26771	Sequence 26771, A
42	57	30.0	53	US-10-057-498-26771	Sequence 26771, A
43	57	30.0	53	US-10-057-498-26771	Sequence 26771, A
44	57	30.0	266	US-10-431-652-6190	Sequence 6190, Ap
45	57	30.0	601	US-10-369-493-11844	Sequence 11844, A

ALIGNMENTS

```
RESULT 1
US-10-425-114-70663
; Sequence 70663, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Zhou, Jindong
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313) B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70663
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73064A07_Flt.pep
US-10-425-114-70663

Query Match          40.3%; Score 76.5; DB 6; Length 245;
Best Local Similarity 47.6%; Pred. No. 3.8;
Matches 20; Conservative 2; Mismatches 7; Indels 13; Gaps 2;

QY      2  RRVVRRVVRVVRVVRVVRVVRVVRVVRVVR 36
Db      6  RWLWR-----RWVRLRRRRRRRRRWVVR 41

RESULT 2
US-10-264-237-2376
; Sequence 2376, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4131P1
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2376
```



```

RESULT 2
PCT-US02-04812-11
; Sequence 11, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PasteoSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-11

```

Query Match	100.0%;	Score 190;	DB 1;	Length 36;
Best Local Similarity	100.0%;	Pred. No. 5.6e-15;		
Matches 36;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 3
US-09-785-058-11
; Sequence 11, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-11

```

Query Match	100.0%	Score 190;	DB 21;	length 36;
Best Local Similarity	100.0%	Pred. No. 5.6e-15;		
Matches	36;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

```

RESULT 4
US-09-785-059-11
Sequence 11, Application US/09785059
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A33577 / 072396, 0217
CURRENT APPLICATION NUMBER: US/09/785, 059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Fastseq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 36
TYPE: PRT

```

```

; ORGANISM: Artificial sequence
;
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-03-785-059-11

```

Query Match	100.0%;	Score 190;	DB 21;	Length 36;
Best Local Similarity	100.0%;	Pred. No. 5.6e-15;		
Matches 36; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 5
US-10-079-075-11
; Sequence 11, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Miezner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 073396,0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ. ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-11

```

Query Match	100.0%	Score 190;	DB 24;	Length 36;
Best Local Similarity	100.0%	Pred. No. 5.6e-15;		
Matches 36;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 6
PCT-US02-04432-12
; Sequence 12, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PC/ 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
PCT-US02-04432-12

```

Query Match	92.6%	Score 176;	DB 1;	Length 48;
Best Local Similarity	100.0%	Pred. No. 3.2e-13;		
Matches 33; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 7

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:55:47 ; Search time 155.872 Seconds

(without alignments)
148.906 Million cell updates/sec

Title: US-10-079-075-11

Perfect score: 190
Sequence: 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/prodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/prodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/prodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/prodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/prodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/prodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/prodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/prodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/prodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/prodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/prodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/prodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/prodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/prodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/prodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/prodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/prodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/prodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/prodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/prodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/prodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	190	100.0	36	1	PCT-US02-04432-11
2	190	100.0	36	1	PCT-US02-04812-11
3	190	100.0	36	21	US-09-785-058-11
4	190	100.0	36	21	US-09-785-059-11
5	190	100.0	36	24	US-10-079-075-11
6	176	92.6	48	1	PCT-US02-04432-12

7	176	92.6	48	1	PCT-US02-04812-12	Sequence 12, Appl
8	176	92.6	48	21	US-09-785-058-12	Sequence 12, Appl
9	176	92.6	48	21	US-09-785-059-12	Sequence 12, Appl
10	176	92.6	48	24	US-10-079-075-12	Sequence 12, Appl
11	176	92.6	36	1	PCT-US02-04432-6	Sequence 6, Appl
12	176	92.6	36	1	PCT-US02-04812-6	Sequence 6, Appl
13	176	92.6	36	21	US-09-785-058-6	Sequence 6, Appl
14	176	92.6	36	21	US-09-785-059-6	Sequence 6, Appl
15	176	92.6	36	24	US-10-079-075-6	Sequence 6, Appl
16	176	92.6	42	1	PCT-US02-04432-7	Sequence 7, Appl
17	176	92.6	42	1	PCT-US02-04812-7	Sequence 7, Appl
18	176	92.6	42	21	US-09-785-058-7	Sequence 7, Appl
19	176	92.6	42	21	US-09-785-059-7	Sequence 7, Appl
20	176	92.6	42	24	US-10-079-075-7	Sequence 7, Appl
21	176	92.6	48	1	PCT-US02-04432-8	Sequence 8, Appl
22	176	92.6	48	1	PCT-US02-04812-8	Sequence 8, Appl
23	176	92.6	48	21	US-09-785-058-8	Sequence 8, Appl
24	176	92.6	48	21	US-09-785-059-8	Sequence 8, Appl
25	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
26	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
27	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
28	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
29	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
30	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
31	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
32	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
33	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
34	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
35	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
36	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
37	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
38	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
39	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
40	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
41	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
42	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
43	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
44	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl
45	176	92.6	48	24	US-10-079-075-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
PCT-US02-04432-11
; Sequence 11, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mielzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04432-11

Query Match 100.0%; Score 190; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR	36
DB	1	VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR	36


```
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-7

Query Match      70.5%; Score 134; DB 9; Length 42;
Best Local Similarity 88.9%; Pred. No. 1.9e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRR 36
DB      7 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRR 42

RESULT 13
US-09-785-059-8
; Sequence 8, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1.
US-09-785-059-8

Query Match      70.5%; Score 134; DB 9; Length 48;
Best Local Similarity 88.9%; Pred. No. 1.9e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRR 36
DB      13 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRR 48

RESULT 14
US-10-079-075-8
; Sequence 8, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-8

Query Match      70.5%; Score 134; DB 9; Length 48;
Best Local Similarity 88.9%; Pred. No. 1.9e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRR 36
DB      13 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRR 48

RESULT 15
US-09-785-058-8
; Sequence 8, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-8

Query Match      70.5%; Score 134; DB 9; Length 48;
Best Local Similarity 88.9%; Pred. No. 1.9e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRR 36
DB      13 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRR 48
```

Search completed: June 9, 2003, 12:34:11
Job time : 19.1489 secs

FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-6

Query Match 70.5%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 1.4e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36
DB 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 8

US-10-079-075-6
Sequence 6, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-6

Query Match 70.5%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 1.4e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36
DB 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 9

US-09-785-058-6
Sequence 6, Application US/09785058
Publication No. US20030036627A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A 34001 / 072396.0222
CURRENT APPLICATION NUMBER: US/09/785.058
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-6

Query Match 70.5%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 1.4e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36
DB 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 10

US-09-785-059-7
Sequence 7, Application US/09785059
Patent No. US20020169279A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 42
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-7

Query Match 70.5%; Score 134; DB 9; Length 42;
Best Local Similarity 88.9%; Pred. No. 1.6e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36
DB 7 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 42

RESULT 11

US-10-079-075-7
Sequence 7, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 42
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-7

Query Match 70.5%; Score 134; DB 9; Length 42;
Best Local Similarity 88.9%; Pred. No. 1.6e-09;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36
DB 7 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRR 42

RESULT 12

US-09-785-058-7
Sequence 7, Application US/09785058
Publication No. US20030036627A1

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-11
```

```
Query Match          100.0%; Score 190; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.2e-16;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
Db 1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
```

```
RESULT 3
US-09-785-058-11
; Sequence 11, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-11
```

```
Query Match          100.0%; Score 190; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.2e-16;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
Db 1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
```

```
RESULT 4
US-09-785-059-12
; Sequence 12, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-059-12
```

```
Query Match          92.6%; Score 176; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 33
Db 13 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 45
```

```
RESULT 5
US-10-079-075-12
; Sequence 12, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-10-079-075-12
```

```
Query Match          92.6%; Score 176; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 33
Db 13 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 45
```

```
RESULT 6
US-09-785-058-12
; Sequence 12, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-058-12
```

```
Query Match          92.6%; Score 176; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 33
Db 13 VRRVRRVVRRVRRVRRVRRVRRVRRVRRVRRVRR 45
```

```
RESULT 7
US-09-785-059-6
; Sequence 6, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 12:01:36 ; Search time 19.1489 Seconds
(without alignments)
194.092 Million cell updates/sec

Title: US-10-079-075-11

Perfect score: 190
Sequence: 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	190	100.0	36	9	US-09-785-059-11
2	190	100.0	36	9	US-10-079-075-11
3	190	100.0	36	9	US-09-785-058-11
4	176	92.6	48	9	US-09-785-059-12
5	176	92.6	48	9	US-10-079-075-12
6	176	92.6	48	9	US-09-785-058-12
7	134	70.5	36	9	US-09-785-059-6
8	134	70.5	36	9	US-10-079-075-6
9	134	70.5	36	9	US-09-785-058-6
10	134	70.5	42	9	US-09-785-059-7
11	134	70.5	42	9	US-10-079-075-7
12	134	70.5	42	9	US-09-785-058-7
13	134	70.5	48	9	US-09-785-059-8
14	134	70.5	48	9	US-10-079-075-8
15	134	70.5	48	9	US-09-785-058-8
16	130	68.4	24	9	US-09-785-059-10
17	130	68.4	24	9	US-10-079-075-10
18	130	68.4	24	9	US-09-785-058-10
19	96	50.5	31	9	US-09-785-059-3

20	96	50.5	31	9	US-10-079-075-3	Sequence 3, Appl1
21	96	50.5	31	9	US-09-785-058-3	Sequence 3, Appl1
22	88	46.3	24	9	US-09-785-059-5	Sequence 5, Appl1
23	88	46.3	24	9	US-10-079-075-5	Sequence 5, Appl1
24	88	46.3	24	9	US-09-785-058-5	Sequence 5, Appl1
25	67	35.3	31	9	US-09-785-059-2	Sequence 2, Appl1
26	67	35.3	31	9	US-10-079-075-2	Sequence 2, Appl1
27	67	35.3	31	9	US-09-785-058-2	Sequence 2, Appl1
28	67	35.3	63	10	US-09-732-665-10	Sequence 10, Appl1
29	67	35.3	63	10	US-09-732-665-8	Sequence 8, Appl1
30	67	35.3	63	10	US-09-732-665-9	Sequence 9, Appl1
31	62	32.6	39	4	US-10-081-816-40	Sequence 40, Appl1
32	61	32.1	12	9	US-09-785-059-9	Sequence 9, Appl1
33	61	32.1	12	9	US-10-079-075-9	Sequence 9, Appl1
34	61	32.1	12	9	US-09-785-058-9	Sequence 9, Appl1
35	59	31.1	76	10	US-09-732-665-6	Sequence 6, Appl1
36	59	31.1	77	9	US-09-992-896-9	Sequence 9, Appl1
37	59	31.1	77	10	US-09-815-666-31	Sequence 31, Appl1
38	59	31.1	83	10	US-09-732-665-7	Sequence 7, Appl1
39	58	30.5	28	9	US-09-785-059-1	Sequence 1, Appl1
40	58	30.5	28	9	US-10-079-075-1	Sequence 1, Appl1
41	58	30.5	28	9	US-09-785-058-1	Sequence 1, Appl1
42	57.5	30.3	66	10	US-09-864-761-33833	Sequence 33833, A
43	54	28.4	31	9	US-09-764-891-4170	Sequence 4170, Ap
44	53	27.9	24	9	US-09-738-626-6003	Sequence 6003, Ap
45	50.5	26.6	19	10	US-09-864-761-38480	Sequence 38480, A

ALIGNMENTS

RESULT 1
US-09-785-059-11
; Sequence 11, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mieczner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-11

Query Match: 100.0%; Score 190; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.2e-16;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 2
US-10-079-075-11
; Sequence 11, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mieczner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12

DE SEN virus protein fragment SEQ ID NO: 122.
XX SEN virus; SENV; gastrointestinal tract disorder; inflammatory disease;
KM proliferative disorder; hepatopathy; hepatitis; viral infection;
KM vaccination; gene therapy.
XX Hepatitis virus.
XX WO200028039-A2.
PN 18-MAY-2000.
PD 09-NOV-1999; 99WO-EP08566.
XX 10-NOV-1998; 98IT-MI02437.
PR 30-APR-1999; 99IT-MI00923.
PR 14-MAY-1999; 99EP-0830298.
PR 16-JUL-1999; 99EP-0113932.
XX (DIAS-) DIASORIN SRL.
PA Primi D, Fiorallisi G, Mantero GL, Mattioli S, Sottini A;
PI Bonelli F, Vaglini L, Olivero P, Dal Corso A, Bonelli M;
DR WPI; 2000-376551/32.
XX Nucleic acids representing the genome of the SEN virus (SENV) and
PT encoded proteins, useful for treatment of hepatopathies, inflammatory
PT diseases and proliferative disorders such as cancer.
XX Claim 1; Page 356-358; 392pp; English.
PS The present invention is concerned with the sequence of the genome of the
CC SEN virus (SENV), and the proteins encoded by it. SENV is thought to be
CC the cause of hepatopathies which are not linked to the presence of the
CC hepatitis A, B and E viruses in man. The genome and proteins of this
CC virus can be used in gene therapy and vaccination against the virus,
CC which also causes disorders of the gastrointestinal tract, including
CC Crohn's disease and lupus erythematosus, inflammatory diseases, and
CC proliferative disorders such as cancer.
XX Sequence 743 AA;
SQ
Query Match 33.9%; Score 64.5; DB 21; Length 743;
Best Local Similarity 52.8%; Pred. No. 6.3;
Matches 19; Conservative 3; Mismatches 7; Indels 7; Gaps 2;
QY 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 36 VRR-RRQRRVRRRFFYGRRRGRMR-----RRYIR 64
RESULT 14
AAB75222
ID AAB75222 standard; Protein; 397 AA.
XX AAB75222;
AC AAB75222;
DT 03-APR-2001 (first entry)
XX Drosophila gustatory receptor GR59D.2 protein sequence.
DE Gustatory receptor; fruit fly; taste; pheromone; semiochemical;
KM crop damage; pest control.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX WO200077208-A2.
PN 21-DEC-2000.
PD 14-JUN-2000; 2000WO-US16211.
PF
XX

PR 14-JUN-1999; 99US-0138668.
PR 10-FEB-2000; 2000US-0181704.
XX (UYVA) UNIV YALE.
PA Carlson PJ, Clyne PJ, Warr CG;
PI WPI; 2001-061873/07.
DR N-PSDB; AAF63761.
XX New isolated nucleic acid molecule encoding Drosophila Gustatory
PT Receptor protein useful for e.g. identification of compounds which may
PT be used for pest management -
XX Claim 12; Page 178-179; 227pp; English.
PS This invention relates to polynucleotide sequences AAF63732 - AAF63777
CC which encode Drosophila gustatory receptor proteins represented by
CC sequences AAB75193 - AAB75238. The invention includes methods for
CC determining gustatory receptor ligands. Also included is a method for
CC modulating the expression of the DNA encoding the receptors. The DNA and
CC protein sequences may be used for the identification of compounds,
CC e.g. pheromones and other semiochemicals, which may be used for pest
CC management. The DNA sequences may also be used for behavioural studies
CC involving gustatory systems in various organisms. Also, the DNA sequences
CC may also be used to track down gustatory receptor genes in insects that
CC damage crops or transmit diseases.
XX Sequence 397 AA;
SQ
Query Match 32.6%; Score 62; DB 22; Length 397;
Best Local Similarity 29.5%; Pred. No. 6.7;
Matches 13; Conservative 11; Mismatches 10; Indels 10; Gaps 2;
QY 3 RWRVRRVVRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 86 RISAVLISLITRWYQSRFRIRINQILALVRDPQVRRGWYRR 129
RESULT 15
ID ABB68280 standard; Protein; 681 AA.
XX ABB68280;
AC ABB68280;
DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 31632.
DE Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX WO200171042-A2.
PN 27-SEP-2001.
PD 23-MAR-2001; 2001WO-US09231.
PF 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL12363.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

PT Novel virus, designated sentinel virus I, associated with cryptogenic,
PT non-A-G hepatitis, and polynucleotides and polypeptides of virus useful
PT for detecting SVI virus and/or SVI virus infection -

XX Example 1; Page 58-59; 65pp; English.

XX The present sequence represents a protein of Sentinel virus I (SVI).
XX SVI polynucleotides are useful for detecting SVI virus. Probes
CC and primers derived from SVI polynucleotide sequences are useful for
CC identifying and isolating new variants of SVI. SVI polynucleotides are
CC useful for detecting SVI virus, producing SVI polypeptides, constructing
CC SVI-based expression/transduction vectors and as antisense
CC oligonucleotides or for construction of antisense SVI vectors. Antisense
CC SVI polynucleotides block expression of SVI proteins and/or SVI viral
CC replication in SVI infected cells, and thus are useful for treating SVI
CC infections. SVI polypeptides are useful in vaccines for preventing SVI
CC infection and for treating SVI infection.

XX Sequence 635 AA;

SO Query Match 35.3%; Score 67; DB 22; Length 635;
Best Local Similarity 61.1%; Pred. No. 2.8;
Matches 22; Conservative 1; Mismatches 9; Indels 4; Gaps 3;

QY 2 RRVRRV-VRVRRVRRVRRVRRVRRVRRVRR 36
| | | | | | | | | | | | | | | | | | | | | |
19 RRRRRLRRRRPRRAVR-RRGRVRR--RRWARR 51

Db

RESULT 9

AAB84457 standard; Protein; 635 AA.

XX AAB84457;

DT 22-AUG-2001 (first entry)

XX Amino acid sequence of a Sentinel virus I (SVI) protein.

XX SVI; viral replication; viral infection; vaccine.

XX Sentinel virus I.

XX WO200142299-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000MO-IB02011.

XX 10-DEC-1999; 99US-0172696.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX Liu J, Bohenzky RA, Lin Y, Chen BP;

XX WPI; 2001-381643/40.

XX Novel virus, designated sentinel virus I, associated with cryptogenic,
PT non-A-G hepatitis, and polynucleotides and polypeptides of virus useful
PT for detecting SVI virus and/or SVI virus infection -

XX Example 1; Page 60-62; 65pp; English.

XX The present sequence represents a protein of Sentinel virus I (SVI).

XX SVI polynucleotides are useful for detecting SVI virus. Probes

CC and primers derived from SVI polynucleotide sequences are useful for

CC identifying and isolating new variants of SVI. SVI polynucleotides are

CC useful for detecting SVI virus, producing SVI polypeptides, constructing

CC SVI-based expression/transduction vectors and as antisense

CC oligonucleotides or for construction of antisense SVI vectors. Antisense

CC SVI polynucleotides block expression of SVI proteins and/or SVI viral

CC replication in SVI infected cells, and thus are useful for treating SVI

CC infections. SVI polypeptides are useful in vaccines for preventing SVI

CC infection and for treating SVI infection.

XX Sequence 635 AA;

SO Query Match 35.3%; Score 67; DB 22; Length 635;
Best Local Similarity 61.1%; Pred. No. 2.8;
Matches 22; Conservative 1; Mismatches 9; Indels 4; Gaps 3;

QY 2 RRVRRV-VRVRRVRRVRRVRRVRRVRRVRR 36
| | | | | | | | | | | | | | | | | | | | | |
19 RRRRRLRRRRPRRAVR-RRGRVRR--RRWARR 51

Db

RESULT 10

AU63025 standard; Protein; 96 AA.

XX AU63025;

DT 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #23921.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; EUSA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001MO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'walsonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59630.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris -

XX Example 1; SEQ ID No 24220; 1069pp; English.

XX Sequences AU9105-AU68017 represent Propionibacterium acnes immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention

CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of P. acnes polypeptides and

CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by

CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

Best Local Similarity 38.2%; Pred. No. 0.64;
Matches 13; Conservative 11; Mismatches 9; Indels 1; Gaps 1;

QY 4 VWR-RVVRVVRVVRVVRVVRVVRVVRVVR 36
DB 488 LWRNRLRLILMLWRNRLILMLILRLRNLRLR 521

RESULT 6

AAB03840
ID AAB03840 standard; protein; 120 AA.

AC AAB03840;

DT 20-OCT-2000 (first entry)

DE Protein fragment #2 used in TT virus antibody determination.

KM Serum type classification; TT virus; antibody; viral infection;
identify; treatment.

OS TT virus.

PN JP2000135087-A.

PD 16-MAY-2000.

PF 29-OCT-1998; 98JP-0309208.

PR 29-OCT-1998; 98JP-0309208.

PA (SRLS-) SRL KK.

DR WPI; 2000-415430/36.

PT Peptides for determination of anti-TT virus antibody and method for
serum classification of TT virus using the peptides -

PS Claim 1; Page 7; 12pp; Japanese.

CC A method for serum type classification of TT virus (also known as
hepatitis TT virus) has been identified. The method relies on the use of
peptide fragments of the virus. The invention also relates to the use of
CC TT virus peptides for anti-TT virus antibody determination. The anti-TT
CC virus antibodies and the serum type classification method, can be used to
screen TT virus, to determine its route of infection, and seroconversion.
CC The classification of TT virus may lead to improved treatment of viral
CC disease. The present sequence represents a fragment of TT virus protein
used in the course of the invention.

SO Sequence 120 AA;

Query Match 37.6%; Score 71.5; DB 21; Length 120;
Best Local Similarity 47.9%; Pred. No. 0.16;

Matches 23; Conservative 1; Mismatches 7; Indels 17; Gaps 4;

QY 2 RRVRVVRVVRVVRVVRVVRVVRVVRVVR 36
DB 11 RRPVR--RRVRVVRVVRVVRVVRVVRVVR--RTVARRRRGRVWR 54

RESULT 7

AAB84458
ID AAB84458 standard; Protein; 634 AA.

AC AAB84458;

DT 22-AUG-2001 (first entry)

DE Amino acid sequence of a Sentinel virus I (SVI) protein.

KM SVI; viral replication; viral infection; vaccine.

OS Sentinel virus I.
XX WO200142299-A2.
PN 14-JUN-2001.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-IB02011.

PR 10-DEC-1999; 99US-0172696.

PA (HOFF) ROCHE DIAGNOSTICS GMBH.

PI Liu J, Bohenzky RA, Lin Y, Chen BP;

DR WPI; 2001-381643/40.

PT Novel virus, designated sentinel virus I, associated with cryptogenic,
non-a-g hepatitis, and polynucleotides and polypeptides of virus useful
for detecting SVI virus and/or SVI virus infection -

PS Example 1; Page 62-64; 65pp; English.

CC The present sequence represents a protein of Sentinel virus I (SVI).
CC SVI polynucleotides are useful for detecting SVI virus. Probes
CC and primers derived from SVI polynucleotide sequences are useful for
CC identifying and isolating new variants of SVI. SVI polynucleotides are
CC useful for detecting SVI virus, producing SVI polypeptides, constructing
CC SVI-based expression/transduction vectors and as antisense
CC oligonucleotides or for construction of antisense SVI vectors. Antisense
CC SVI polynucleotides block expression of SVI proteins and/or SVI viral
CC replication in SVI infected cells, and thus are useful for treating SVI
CC infections. SVI polypeptides are useful in vaccines for preventing SVI
CC infection and for treating SVI infection.

SO Sequence 634 AA;

Query Match 35.3%; Score 67; DB 22; Length 634;
Best Local Similarity 61.1%; Pred. No. 2.8;

Matches 22; Conservative 1; Mismatches 9; Indels 4; Gaps 3;

QY 2 RRVVRV-VRVVRVVRVVRVVRVVRVVRVVR 36
DB 19 RRRVRLRRRPRRAVR-RRGRVVR--RRVARR 51

RESULT 8
AAB84456
ID AAB84456 standard; Protein; 635 AA.

AC AAB84456;

DT 22-AUG-2001 (first entry)

DE Amino acid sequence of a Sentinel virus I (SVI) protein.

KM SVI; viral replication; viral infection; vaccine.

OS Sentinel virus I.

PN WO200142299-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-IB02011.

PR 10-DEC-1999; 99US-0172696.

PA (HOFF) ROCHE DIAGNOSTICS GMBH.

PI Liu J, Bohenzky RA, Lin Y, Chen BP;

DR WPI; 2001-381643/40.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145324.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:35:06 ; Search time 38.8085 Seconds
(without alignments)
123.607 Million cell updates/sec

Title: US-10-079-075-11

Perfect score: 190

Sequence: 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A-Geneseq, 101002.*

1: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneeq/geneeq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	80.5	42.4	133	22	ABB61318	Drosophila melanog
2	73	38.4	82	21	AA628909	Arabidopsis thalia
3	73	38.4	95	21	AA630720	Arabidopsis thalia
4	72.5	38.2	136	23	ABB30000	Human polypeptide
5	72.5	38.2	643	22	ABG22551	Novel human diagno
6	71.5	37.6	120	21	AA503840	Protein fragment #
7	67	35.3	634	22	AA844458	Amino acid sequenc
8	67	35.3	635	22	AA844456	Amino acid sequenc
9	67	35.3	635	22	AA844457	Amino acid sequenc
10	66.5	35.0	96	22	AAU63025	Protonibacterium

11	64.5	33.9	40	16	AA844926	Alpha-helix-formin
12	64.5	33.9	105	21	AA654778	Arabidopsis thalia
13	64.5	33.9	743	21	AA611535	SEN virus protein
14	62	32.6	397	22	AA675222	Drosophila gustato
15	62	32.6	681	22	AA668280	Drosophila melanog
16	61	32.1	77	22	AAU53862	Protonibacterium
17	61	32.1	898	18	AA631853	Mycobacterium tube
18	60.5	31.8	147	21	AA645794	Arabidopsis thalia
19	60.5	31.8	727	23	AA647989	Simian TTV CH71 SE
20	59.5	31.3	526	22	AA618314	Novel human diagno
21	59.5	31.3	758	21	AA611540	SEN virus protein
22	59.5	31.3	765	23	AA647987	Simian TTV CH65-1
23	59	31.1	120	21	AA603839	Protein fragment #
24	59	31.1	142	21	AA641902	Arabidopsis thalia
25	59	31.1	261	21	AA641901	Arabidopsis thalia
26	59	31.1	280	21	AA641900	Arabidopsis thalia
27	59	31.1	756	23	AA647795	TT virus clone X94
28	59	31.1	768	22	AA644454	Amino acid sequenc
29	59	31.1	770	20	AA699082	Non-B, non-C, non-
30	59	31.1	770	21	AA697179	TT virus (TTV-GH1)
31	59	31.1	770	22	AA637535	TT virus ORF1 prot
32	59	31.1	831	22	AA684455	Amino acid sequenc
33	58	30.5	39	18	AA606684	Proctamine-like pep
34	58	30.5	66	22	AAU51034	Protonibacterium
35	58	30.5	762	21	AA611546	SEN virus protein
36	57.5	30.3	66	22	AA627894	Human peptide #545
37	57.5	30.3	66	22	AA633065	Peptide #571 encod
38	57.5	30.3	66	22	AA618535	Protein #534 expres
39	57.5	30.3	66	22	AA653863	Human brain expres
40	57.5	30.3	66	22	AA666250	Human bone marrow
41	57.5	30.3	66	22	AA614120	Peptide #554 encod
42	57.5	30.3	66	22	AA626527	Peptide #564 encod
43	57.5	30.3	66	22	AA601858	Peptide #540 encod
44	57.5	30.3	66	23	ABG35897	Human peptide enco
45	57	30.0	53	22	AAU59076	Protonibacterium

ALIGNMENTS

RESULT 1
ID ABB61318 standard; Protein; 133 AA.
AC ABB61318;
XX
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 10746.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX
XX W0200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US029231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PMD, Myers EW;
PI
XX MPI; 2001-65660/75.
XX
XX N-PSDB; ABL05421.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE ORF1.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OC NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JT03F;
 RA Okamoto H.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JT03F;
 RX MEDLINE=21844401; PubMed=11855633;
 RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,
 RA Okamoto H.;
 RT "Analysis of the complete genomes of thirteen TT virus variants
 RT classifiable into the fourth and fifth genetic groups, isolated from
 RT viremic infants.";
 RT Arch. Virol. 147:21-41(2002).
 DR EMBL; AB064599; BAB79326.1; -.
 DR InterPro; IPR004219; TVirus_Unk.
 DR Pfam; PF02956; TT_ORF1.1.
 SQ SEQUENCE 746 AA; 89191 MW; PB516A5190D61DC7 CRC64;

Query Match 35.0%; Score 66.5; DB 12; Length 746;
 Best Local Similarity 52.6%; Pred. No. 3.3;
 Matches 20; Conservative 2; Mismatches 9; Indels 7; Gaps 3;

QY 2 RRVRRVVVRRVRRV---VRRVRRVVRRVVRRVRR 36
 DB 8 RRRWR--RRWGRLRRRRRTFR--RRRRVSR 41

RESULT 14
 OSV7C1 PRELIMINARY; PRT; 49 AA.
 AC OSV7C1;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE ORF1 (Fragment).
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OC NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21844401; PubMed=11855633;
 RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,
 RA Okamoto H.;
 RT "Analysis of the complete genomes of thirteen TT virus variants
 RT classifiable into the fourth and fifth genetic groups, isolated from
 RT viremic infants.";
 RT Arch. Virol. 147:21-41(2002).
 DR EMBL; AB064629; BAB79402.1; -.
 FT NON_TER
 SQ SEQUENCE 49 AA; 7034 MW; C1EA6BBF6A1DCFD CRC64;

Query Match 34.7%; Score 66; DB 12; Length 49;
 Best Local Similarity 57.1%; Pred. No. 0.29;
 Matches 20; Conservative 1; Mismatches 12; Indels 2; Gaps 2;

QY 2 RRVRRVVVRRVRRV---VRRVRRVVRRVVRRVRR 35
 DB 15 RRRWR--RRWGRLRRRRRTFR--RRRRVSR 48

RESULT 15
 OSV7G0 PRELIMINARY; PRT; 744 AA.
 AC OSV7G0;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE ORF1.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OC NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT44F;
 RA Okamoto H.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT44F;
 RX MEDLINE=21844401; PubMed=11855633;
 RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,
 RA Okamoto H.;
 RT "Analysis of the complete genomes of thirteen TT virus variants
 RT classifiable into the fourth and fifth genetic groups, isolated from
 RT viremic infants.";
 RT Arch. Virol. 147:21-41(2002).
 DR EMBL; AB064605; BAB79350.1; -.
 DR InterPro; IPR004219; TVirus_Unk.
 DR Pfam; PF02956; TT_ORF1.1.
 SQ SEQUENCE 744 AA; 87602 MW; 33C2D5349CFEBA69 CRC64;

QY 2 RRVRRVVVRRVRRV---VRRVRRVVRRVVRRVRR 35
 DB 15 RRRWR--RRWGRLRRRRRTFR--RRRRVSR 48

Search completed: June 9, 2003, 12:01:11
 Job time : 34.0638 secs


```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PT-TTV6;
RA Okamoto H.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PT-TTV6;
RX MEDLINE=20534983; PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RT Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
RT phylogenetic relatedness.";
RT Virology 277:368-378(2000).
RL EMBL; AB041957; BAB19308.1; -
DR InterPro; IPR004219; TTVirus_Unk.
DR Pfam; PF02956; TT_ORF1.1.
SQ SEQUENCE 735 AA; 86132 MW; 9ED818D6BE6FA5D3 CRC64;

Query Match 37.4%; Score 71; DB 12; Length 735;
Best Local Similarity 57.6%; Pred. No. 0.98;
Matches 19; Conservative 1; Mismatches 11; Indels 2; Gaps 1;

QY 2 RRVRRVVR--VVRVRRVRRVRRVRRVRR 32
Db 25 RRTWRRRPRRAVRRVRRRGGRRLYRRYRR 57

RESULT 6
Q9DUB7 PRELIMINARY; PRT; 720 AA.
AC Q9DUB7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN RP SEQUENCE FROM N.A.
RC STRAIN=AT-TTV3;
RA Okamoto H.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AT-TTV3;
RX MEDLINE=20534983; PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RT Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
RT phylogenetic relatedness.";
RT Virology 277:368-378(2000).
RL EMBL; AB041957; BAB19308.1; -
DR InterPro; IPR004219; TTVirus_Unk.
DR Pfam; PF02956; TT_ORF1.1.
SQ SEQUENCE 720 AA; 85948 MW; 4915FE9AE503E15 CRC64;

Query Match 36.8%; Score 70; DB 12; Length 720;
Best Local Similarity 54.5%; Pred. No. 1.3;
Matches 18; Conservative 0; Mismatches 7; Indels 8; Gaps 2;

QY 4 VWRVRRVVRVRRVRRVRRVRRVRRVRR 36
Db 2 VWR-----RWRRRRRRVRRVRRVRR 26

RESULT 7
Q9IFV0 PRELIMINARY; PRT; 766 AA.
AC Q9IFV0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE
```

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative capsid protein.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN RP SEQUENCE FROM N.A.
RC STRAIN=PMV;
RX MEDLINE=20409089; PubMed=10950985;
RA Hallatt R.L., Clewley J.P., Bobet F., McKiernan P.J., Teo C.G.;
RT "Characterization of a highly divergent TT virus genome.";
RT J. Gen. Virol. 81:2273-2279(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PMV;
RA Hallatt R.L., Clewley J.P., Teo C.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF261761; AAF82559.1; -
DR InterPro; IPR004219; TTVirus_Unk.
DR Pfam; PF02956; TT_ORF1.1.
SQ SEQUENCE 766 AA; 90844 MW; 22E5A3BBA3CC8187 CRC64;

Query Match 36.8%; Score 70; DB 12; Length 766;
Best Local Similarity 41.3%; Pred. No. 1.3;
Matches 19; Conservative 0; Mismatches 13; Indels 14; Gaps 1;

QY 5 WRRVRRVRRVRRVRRVRRVRRVRRVRR 36
Db 3 WRRVRRRRRRVRRRRRTTWRRRPRPRRRRTAATRRRGVRRVRR 48

RESULT 8
Q9DT81 PRELIMINARY; PRT; 748 AA.
AC Q9DT81;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN RP SEQUENCE FROM N.A.
RC STRAIN=TYM9;
RX MEDLINE=20568739; PubMed=1118348;
RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
RT Sai T., Sugai Y.;
RT "TT virus mRNAs detected in the bone marrow cells from an infected
RT individual.";
RT Biochem. Biophys. Res. Commun. 279:700-707(2000).
RL EMBL; AB050446; BAB19928.1; -
DR InterPro; IPR004219; TTVirus_Unk.
DR Pfam; PF02956; TT_ORF1.1.
SQ SEQUENCE 748 AA; 88552 MW; D65CCB2CA5CE26F CRC64;

Query Match 36.1%; Score 68.5; DB 12; Length 748;
Best Local Similarity 55.6%; Pred. No. 1.9;
Matches 20; Conservative 3; Mismatches 10; Indels 3; Gaps 2;

QY 2 RRVRRV--VVRVRRVRRVRRVRRVRRVRR 36
Db 18 RRRVRLRTRRRPRRLVRRRRKRYR--VRRRRWGRR 51

RESULT 9
Q9ID04 PRELIMINARY; PRT; 750 AA.
AC Q9ID04;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE
```


FT PROPEP 1 24 BY SIMILARITY
 FT CHAIN 25 193 MAJOR CORE PROTEIN.
 FT SITE 24 25 CLEAVAGE (BY ADENOVIRUS PROTEASE)
 FT (POTENTIAL).
 SQ SEQUENCE 193 AA; 21358 MM; 43137E07DB379DD0 CRC64;

Query Match 26.8%; Score 51; DB 1; Length 193;
 Best Local Similarity 40.5%; Pred. No. 4.7;
 Matches 15; Conservative 5; Mismatches 11; Indels 6; Gaps 2;

QY 2 RRVVRR-----VVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 32
 Db 103 RRIARRRSTTAMRAAALRLRRARRTGRRAVRLRAARR 139

Search completed: June 9, 2003, 11:56:42
 Job time : 7.2766 secs

FT STRAND CLEAVAGE AND REJOINING
(BY SIMILARITY)
SQ SEQUENCE 305 AA; 34277 MW; 48C03BD3A4A9420F CRC64;
Query Match 28.9%; Score 55; DB 1; Length 305;
Best local Similarity 36.2%; Pred. No. 2.3;
Matches 17; Conservative 7; Mismatches 11; Indels 12; Gaps 3;
QY 2 RRVRRVVRVVR---W-VRRVRVVRVVR-----VRRVRR 36
196 RHSMWNVVRIIRRGHWTVERRLRRAVRLVREKLAEPELLARSLRR 242
RESULT 11
ERMS_STRFR STANDARD; PRT; 319 AA.
ID ERMS_STRFR STANDARD; PRT; 319 AA.
AC P45439;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA adenine N-6-methyltransferase (EC 2.1.1.48) (Macrolide-
lincomamide-streptogramin B resistance protein) (Erythromycin
resistance protein).
GN ERMSF OR TUBA.
OS Streptomyces fradiae.
OC Bacteria; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
ON NCBI_TaxID=1906;
RX STRAIN=FROM N.A.
RX STRAIN=NRRL 2338;
RX MEDLINE=88169508; PubMed=3127381;
RX Kaminaka S., Weisblum B.;
RT "Translational attenuation control of ermSF, an inducible resistance
determinant encoding RNA N-methyltransferase from Streptomyces
fradiae.";
RT J. Bacteriol. 170:1800-1811(1988).
CC -1- FUNCTION: THIS PROTEIN PRODUCES A DIMETHYLATION OF THE ADENINE
CC RESIDUE AT POSITION 2058 IN 23S RNA, RESULTING IN REDUCED
CC AFFINITY BETWEEN RIBOSOMES & MACROLIDE-LINCOMAMIDE-STREPTOGRAMIN B
CC ANTIOTICS.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing N(6)-methyladenine.
CC -1- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19269; AAA26742.1; -
DR InterPro; IPR001737; RNA_A_dimeth.
DR InterPro; IPR000051; SAM_Bind.
DR Pfam; PF00398; RnaAD; 1.
DR PROSITE; PS01131; RNA_A_DIMETH; 1.
KW Antibiotic resistance; Transferase; Methyltransferase.
SQ SEQUENCE 319 AA; 35527 MW; 3A543FA22CBB7DB CRC64;
Query Match 28.9%; Score 55; DB 1; Length 319;
Best local Similarity 45.2%; Pred. No. 2.5;
Matches 14; Conservative 4; Mismatches 13; Indels 0; Gaps 0;
QY 2 RRVRRVVRVVRVVRVVRVVRVVRVVRVVR 32
208 RHEWRLIGVRSRRFPVPRVDSGILRIERR 238
RESULT 12
WECA_YERPE

ID WECA_YERPE STANDARD; PRT; 365 AA.
AC Q8ZAE1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase (EC
DE 2.4.1.-) (UDP-GlcNAc:undecaprenyl-phosphate GlcNAc-1-phosphate
DE transferase).
GN WECA OR RFE OR YP03866.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
ON NCBI_TaxID=632;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titchell R.W., Holden M.T.G.,
RA Prentice M.B., Sebatina M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.,
RL "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
CC -1- FUNCTION: Catalyzes the synthesis of Und-PP-GlcNAc (lipid I), the
CC first lipid-linked intermediate involved in ECA synthesis. This
CC lipid is also an acceptor for the addition of subsequent sugars to
CC complete the biosynthesis of O-antigen (by similarity).
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + undecaprenyl
CC monophosphate = UDP + undecaprenyl N-acetyl-alpha-D-glucosaminyl
CC pyrophosphate.
CC -1- COFACTOR: Magnesium and manganese (By similarity).
CC -1- PATHWAY: Synthesis of enterobacterial common antigen (ECA).
CC Synthesis of lipopolysaccharide O-antigen.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. WECA
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ414159; CAC93334.1; -
DR InterPro; IPR000715; Glycos_transf_4; 1_
DR Pfam; PF00953; Glycos_transf_4; 1_
KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase;
KW Transmembrane; Inner membrane; Complete proteome.
KW TRANSMEM 3
FT TRANSMEM 3
FT TRANSMEM 45
FT TRANSMEM 69
FT TRANSMEM 89
FT TRANSMEM 99
FT TRANSMEM 119
FT TRANSMEM 122
FT TRANSMEM 132
FT TRANSMEM 157
FT TRANSMEM 177
FT TRANSMEM 187
FT TRANSMEM 207
FT TRANSMEM 213
FT TRANSMEM 233
FT TRANSMEM 242
FT TRANSMEM 262
FT TRANSMEM 293
FT TRANSMEM 315
FT TRANSMEM 335
FT TRANSMEM 365
SQ SEQUENCE 365 AA; 40766 MW; DD9EAB886208B397F CRC64;
Query Match 28.2%; Score 53.5; DB 1; Length 365;
Best local Similarity 47.6%; Pred. No. 4.4;
Matches 10; Conservative 6; Mismatches 4; Indels 1; Gaps 1;
QY 1 VRRVRRVVRVVRVVRVVRVVR 21

```

DR Pfam; PF00512; signal; 1.
DR Pfam; PF00989; PAS; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR Prodom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00073; HPT; 2.
DR SMART; SM00388; HATK; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00448; REC; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50113; PAC; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
DR Sensory transduction; Transferrase; Kinase; Phosphorylation;
KM Transmembrane; Inner membrane; Transcription regulation; 3D-structure;
KM Complete proteome.
FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 26 46 POTENTIAL.
FT DOMAIN 47 57 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 POTENTIAL.
FT DOMAIN 79 778 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 153 223 PAS.
FT DOMAIN 226 278 PAC.
FT DOMAIN 289 507 HISTIDINE KINASE.
FT DOMAIN 527 643 RESPONSE_REGULATORY.
FT DOMAIN 644 778 SECONDARY_TRANSMITTER_DOMAIN (POTENTIAL).
FT MOD_RES 292 292 PHOSPHORYLATION (AUTO-).
FT MOD_RES 576 576 PHOSPHORYLATION (PROBABLE).
FT MOD_RES 717 717 PHOSPHORYLATION (PROBABLE).
FT MUTAGEN 292 292 H->Q: LOSS OF ACTIVITY.
FT MUTAGEN 576 576 D->A: LOSS OF ACTIVITY.
FT MUTAGEN 717 717 H->Q: LOSS OF ACTIVITY.
FT CONFLICT 469 470 MISSING (IN REF. 2).
SQ SEQUENCE 778 AA; 87982 MW; DD61EA6ECF95AD30 CRC64;

Query Match 29.5%; Score 56; DB 1; Length 778;
Best Local Similarity 32.0%; Pred. No. 4.7;
Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 12 VRRVRRVRRVRRVRRVRRVRR 36
Db 750 VGEIEMKEKEMRDRVLAAMAK 774

RESULT 9
SECE_SUISO
ID SECE_SUISO STANDARD; PRT; 62 AA.
AC P58191;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteolysin translocase SecE subunit (Protein transport protein SEC61
DE gamma subunit homolog).
GN SECE OR SSOS5663.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OC NCBI_TaxId=2287;
OX NCBI_TaxId=2287;
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=2133296; PubMed=1142726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Moore M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Erasuo G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Noc H.P., Redder P., Schenk M.B., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.",
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -1- FUNCTION: Involved in protein export (By similarity).

```

```

CC -1- SUBUNIT: Component of the protein translocase complex (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Tail-anchored membrane protein (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE SECE/SEC61-GAMMA FAMILY.
CC -1- SIMILARITY: BELONGS TO THE SECE/SEC61-GAMMA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE00668; AAK40677.1; -.
CC InterPro; IPR001901; SecE.
CC InterPro; IPR004795; SecE_euk_arch.
CC TIGRFAMs; TIGR00327; SecE_euk_arch; 1.
CC PROSITE; PS01067; SECE_SEC61G; FALSE_NEG.
CC Protein transport; Translocation; Transmembrane; Complete proteome.
FT TRANSMEM 40 60 POTENTIAL.
SQ SEQUENCE 62 AA; 7184 MW; CD51874FA24BE9FA CRC64;

Query Match 28.9%; Score 55; DB 1; Length 62;
Best Local Similarity 44.4%; Pred. No. 0.44;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 15 VRRVRRVRRVRRVRRVRRVRR 32
Db 9 WFRRLRDMWRLLITVARK 26

RESULT 10
Y4CG_RHISN
ID Y4CG_RHISN STANDARD; PRT; 305 AA.
AC P55389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable DNA-invertase Y4CG.
GN Y4CG.
OS Rhizobium sp. (strain NGR234).
OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxId=394;
RN NCBI_TaxId=394;
RN SEQUENCE FROM N.A.
RP MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.",
RT Nature 387:394-401(1997).
CC -1- SIMILARITY: BELONGS TO THE SITE-SPECIFIC RECOMBINASE RESOLVASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000068; AAB92422.1; -.
CC HSSP; P03012; ZRSU.
CC InterPro; IPR001822; Recombinase.
CC Pfam; PF00239; resolvase; 1.
CC PROSITE; PS00397; RECOMBINASES_1; 1.
CC PROSITE; PS00398; RECOMBINASES_2; 1.
CC Hypothetical protein; DNA recombination; DNA integration; DNA-binding;
CC DNA invertase; Plasmid.
FT ACT_SITE 23 23 TRANSIENT COVALENT LINKAGE TO DNA DURING

```



```

FT TRANSMEM 2 24 POTENTIAL.
FT TRANSMEM 39 60 POTENTIAL.
FT TRANSMEM 67 84 POTENTIAL.
FT TRANSMEM 126 148 POTENTIAL.
FT TRANSMEM 155 172 POTENTIAL.
FT TRANSMEM 182 204 POTENTIAL.
FT TRANSMEM 209 231 POTENTIAL.
FT TRANSMEM 241 260 POTENTIAL.
FT TRANSMEM 290 312 POTENTIAL.
FT TRANSMEM 317 339 POTENTIAL.
SQ SEQUENCE 357 AA; 40285 MW; B82A186EA21EBD5C CRC64;

Query Match 29.7%; Score 56.5; DB 1; Length 357;
Best Local Similarity 38.5%; Pred. No. 1.8;
Matches 10; Conservative 6; Mismatches 5; Indels 5; Gaps 1;

QY 1 VRRVRRVVRVRRVRRVRRVRRV 26
DB 333 ITRAMR-----ITRWIRRRRRRAKRI 353

RESULT 6
KEGI_ECOLI STANDARD; PRT; 89 AA.
ID KEGI_ECOLI
AC O5282;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative protein Kleg.
GN Kleg.
OS Escherichia coli.
OG Plasmid Inc-beta R751.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / C600;
RX MEDLINE=95291464; PubMed=7773415;
RA Thomas C.M., Smith C.A., Ibbotson J.P., Johnston L., Wang N.;
RT "Evolution of the kora-ori segment of promiscuous Incp plasmids.";
RL Microbiology 141:1201-1210(1995).
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67194; AAC64423.1; -
DR K12 Plasmid; Hypothetical protein.
SQ SEQUENCE 89 AA; 9865 MW; 2986AB96E7051C01 CRC64;

Query Match 29.5%; Score 56; DB 1; Length 89;
Best Local Similarity 38.2%; Pred. No. 0.48;
Matches 13; Conservative 4; Mismatches 11; Indels 6; Gaps 1;

QY 2 RRVRRVVRVRRVRRVRRVRRVRRV 35
DB 37 RRRWRPAGRVLRPRTGTGLMR-----ARWIR 64

RESULT 7
ARCB_ECO57 STANDARD; PRT; 778 AA.
ID ARCB_ECO57
AC P58363;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aerobic respiration control sensor protein archB (EC 2.7.3.-).
GN ARCB OR 24574 OR ECS4089.

```

```

OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxId=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grodeck E.J., Davis N.W., Lim A., Dimantanta E.T., Potamoultis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohnishi E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hatori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC
CC -1- FUNCTION: Member of the two-component regulatory system archB/archC.
CC Sensor-regulator protein for anaerobic repression of the arch
CC module. Activates archA via a four-step phosphorylation. Arch can
CC also dephosphorylate archA by a reverse phosphorylation involving His-
CC 717 and Asp-576 (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -1- PTM: Activation requires a sequential transfer of a phosphate
CC group from a His in the primary transmitter domain, to a Asp in
CC the receiver domain and to a His in the secondary transmitter
CC domain (By similarity).
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RESPONSE REGULATOR DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB005549; BA58344.1; -
DR EMBL; AP002564; BAB37512.1; -
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004359; HIS_KIN_sig.
DR InterPro; IPR003661; HIS_KIN.
DR InterPro; IPR002570; Hpt.
DR InterPro; IPR000700; PAS-assoC.
DR InterPro; IPR000014; PAS_domain.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00072; response_reg.1.
DR Pfam; PF00512; signal.1.
DR Pfam; PF00989; PAS.1.
DR Pfam; PF02518; HATPase_C.1.
DR SMART; SM00387; HATPase_C.1.
DR SMART; SM00073; HPT.1.
DR SMART; SM00073; HPT.1.
DR SMART; SM00388; HSKA.1.
DR SMART; SM00091; PAS.1.
DR SMART; SM00448; REC.1.
DR TIGRFAMs; TIGR00229; sensory_box.1.
DR PROSITE; PS50109; HIS_KIN.1.
DR PROSITE; PS50113; PAC.1.
DR PROSITE; PS50112; PAS.1.

```

DE 50S ribosomal protein L31e.
GN RPL31E OR ABE1087.
OC Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxId=5636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Aoki A., Kousgi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamitsu M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.",
RL DNA Res 6:83-101(1999).
CC -1- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000060; BAA80072.1; -
DR InterPro: IPR000054; Ribosomal_L31e.
DR Pfam: PF01198; Ribosomal_L31e; 1.
DR PROSITE: PS01144; RIBOSOMAL_L31E; FALSE NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 105 AA; 12527 MW; 755DF7999E74A098 CRC64;
QY Query Match 30.5%; Score 58; DB 1; Length 105;
Best Local Similarity 51.5%; Pred. No. 0.32;
Matches 17; Conservative 3; Mismatches 11; Indels 2; Gaps 2;
Db 5 WRRVVRVVR-RWVRVRVVRVVRVVRVVRVVR 36
8 WYVVVLRVYVWGRRTTIRA-IRAVRVVRFVRR 39
RESULT 4
CARB_STRTH STANDARD; PRT; 299 AA.
AC P13079;
DT 01-JUN-1990 (Rel. 13, Created)
DT 01-JUN-1990 (Rel. 13, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE rRNA methyltransferase (EC 2.1.1.-) (Carbomycin-resistance protein).
GN CARB.
OS Streptomyces thermotolerans.
OG Plasmid pO159.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=80858;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11416;
RX MEDLINE=8724811; PubMed=3036668;
RA Epp J.K., Burgelt S.G., Schoner B.E.;
RT "Cloning and nucleotide sequence of a carbomycin-resistance gene from
RT Streptomyces thermotolerans".
RL Gene 53:73-83(1987).
CC -1- FUNCTION: PROBABLE RNA METHYLASE. CARB CONFERS RESISTANCE TO
CC CARBOMYCIN AND SEVERAL OTHER MACROLIDES, LINCOMYCIN AND
CC VANAMYCIN B, BUT NOT TO ALL MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN
CC B ANTIBIOTICS.
CC -1- INDUCTION: THE EXPRESSION OF CARB IS INDUCIBLE BY CERTAIN
CC MACROLIDE ANTIBIOTICS.

CC -1- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M16503; AAC32026.1; -
DR PIR: A26512; A26512.
DR InterPro: IPR001737; RNA_A dimeth.
DR InterPro: IPR000051; SAM_Bind.
DR Pfam: PF00398; RrmAD.1.
DR PROSITE: PS01131; RNA_A_DIMETH.1.
KW Antibiotic resistance; Transferase; Methyltransferase; Plasmid.
SQ SEQUENCE 299 AA; 34138 MW; 63990AA894C044C6 CRC64;
QY Query Match 30.0%; Score 57; DB 1; Length 299;
Best Local Similarity 46.4%; Pred. No. 1.3;
Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
Db 5 WRRVVRVVRVVRVVRVVRVVRVVRVVR 32
198 WRRGRISRMRPVPVAVDSAVLRERR 225
RESULT 5
RFE_PASMTU STANDARD; PRT; 357 AA.
ID RFE_PASMTU
AC Q9CNG8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase
DE (EC 2.4.1.-).
GN RFE OR PM0463.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxId=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + undecaprenyl
CC monophosphate = UMP + undecaprenyl N-acetyl-alpha-D-glucosaminyl
CC pyrophosphate.
CC -1- COFACTOR: Magnesium and manganese (By similarity).
CC -1- PATHWAY: Lipopolysaccharide biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. WECA
CC SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB006082; AAK02547.1; -
DR InterPro: IPR000715; Glycosyltransferase_4.
DR Pfam: PF00953; Glycosyltransferase_4.
KW Lipopolysaccharide biosynthesis; Glycosyltransferase; Transferase;
KW Transmembrane; Inner membrane; Complete proteome.

```

RN [2] IDENTIFICATION, AND TISSUE SPECIFICITY.
RX MEDLINE=20175760; PubMed=10710311;
RA Clyne P.U., Warr C.G., Carlson J.R.;
RT "Candidate taste receptors in Drosophila.";
RL Science 287:1830-1834(2000).
RN [3]
RP IDENTIFICATION.
RX MEDLINE=21407712; PubMed=11516643;
RA Dunipace L., Meister S., McNeely C., Amrein H.;
RT "Spatially restricted expression of candidate taste receptors in the
RT Drosophila gustatory system.";
RL Curr. Biol. 11:822-835(2001).
RN [4]
RP CONCEPTUAL TRANSLATION.
RA Robertson H.;
RL Unpublished observations (NOV-2001).
CC -1- FUNCTION: Probable role in the gustatory response.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Expressed in the adult labellar chemosensory
CC neurons.
CC -1- SIMILARITY: BELONGS TO FAMILY DR-TR OF G-PROTEIN COUPLED
CC RECEPTORS. SUBFAMILY VI.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC gene model prediction.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensing@isb-sib.ch).
-----
DR EMBL; AE003459; AAF46958.2; ALT_SEQ.
DR FLYBASE; FBgn0041235; Gr59c.
KM Hypothetical protein; Receptor; G-protein coupled receptor;
KW transmembrane; Glycoprotein; Multigene family.
FT DOMAIN 1 3
FT TRANSSEM 4 24
FT DOMAIN 25 39
FT TRANSSEM 40 60
FT DOMAIN 61 75
FT TRANSSEM 76 96
FT DOMAIN 97 166
FT TRANSSEM 167 187
FT DOMAIN 188 259
FT TRANSSEM 260 280
FT DOMAIN 281 284
FT TRANSSEM 285 305
FT DOMAIN 306 372
FT TRANSSEM 373 393
FT DOMAIN 394 397
FT CARBOHD 61
FT SEQUENCE 397 AA; 46164 MW; 60CEFF8AC46F9D0E CRC64;
-----
Qy Query Match 32.6%; Score 62; DB 1; Length 397;
Db Best Local Similarity 29.5%; Pred. No. 0.4;
Matches 13; Conservative 11; Mismatches 10; Indels 10; Gaps 2;
-----
Qy 3 RWRRVVRVVRVVRVVRV-RVWRVVRVVRV-----RVVR 36
Db 86 RISAVILSLIRWYQSRFRIRIMQDILAVDRPQVGRWYR 129
-----
RESULT 2
1278 MYCTU STANDARD; PRT; 957 AA.
AC P56877;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical PE-PGRS family protein RV0278c precursor.

```

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:40:31 ; Search time 7.2766 Seconds

(without alignments)
205.199 Million cell updates/sec

Title: US-10-079-075-11

Sequence: 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	32.6	397	1	G59C DROME
2	61	32.1	957	1	Y278 MYCTU
3	58	30.5	105	1	RL31 AERPE
4	57	30.0	299	1	CAR8 STRTH
5	56.5	29.7	357	1	RFE_PASNU
6	56	29.5	89	1	KEG1 ECOLI
7	56	29.5	778	1	ARCB ECOS7
8	56	29.5	778	1	ARCB ECOLI
9	55	28.9	62	1	SECE_SUISO
10	55	28.9	305	1	Y4CG_RHISN
11	55	28.9	319	1	ERMS_STRFR
12	53.5	28.2	365	1	MECA_YERPE
13	53	27.9	256	1	YTS7 MYCTU
14	53	27.9	477	1	Y006 CAEEL
15	51	26.8	193	1	VC07_ADE04
16	51	26.8	354	1	YCUF_HAEIN
17	51	26.8	1582	1	YU30_RALSO
18	51	26.8	4568	1	DYHB_CHIRE
19	50	26.3	85	1	Y02A_BP74
20	49.5	26.1	30	1	PR71_CLUPA
21	49.5	26.1	263	1	ONCM_MOOSE
22	49.5	26.1	1696	1	PCK5_BRACL
23	49	25.8	100	1	HSP2_ALOSE
24	49	25.8	493	1	HXKL_ARATH
25	49	25.8	495	1	GLPK_TREBO
26	49	25.8	498	1	VNUC_IANS
27	49	25.8	498	1	VNUC_IANS
28	49	25.8	498	1	VNUC_IANS
29	49	25.8	525	1	NCAP_MEASH
30	48.5	25.5	31	1	PR72_CLUPA
31	48.5	25.5	31	1	URK_MYCN
32	48	25.0	3511	1	MT15_MOUSE
33	47.5	25.0	102	1	HSP2_PANTR

34	47.5	25.0	314	1	M1A MYCTU
35	47.5	25.0	355	1	RFE_HAEIN
36	47.5	25.0	1146	1	MLC MYCTU
37	47	24.7	72	1	VXIS_BP434
38	47	24.7	72	1	VXIS_LAMB
39	47	24.7	87	1	TAT CAEVG
40	47	24.7	105	1	NDCl_BRAJA
41	47	24.7	204	1	RL15_ARATH
42	47	24.7	211	1	URK_BACHD
43	47	24.7	343	1	Y1S5_SHISO
44	47	24.7	1062	1	YAL3_HUMAN
45	47	24.7	1150	1	IRRI_YEAST

ALIGNMENTS

RESULT 1
ID G59C_DROME STANDARD; PRT; 397 AA.
AC QPM105;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Putative gustatory receptor 59c.
GN GRS9C OR GRS9D.2 OR CG13543 OR CG30186.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G.G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertan P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport I.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Harris A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mewlowski G., Mialshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusslein D.R., Paclet J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT The genome sequence of Drosophila melanogaster.
RL Science 287:2185-2195 (2000).

F:574/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 29.5%; Score 56; DB 1; Length 776;

Best Local Similarity 32.0%; Pred. No. 25;

Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 12 VRRVRRVRRVRRVRRVRRVRR 36
DB 748 VGEWIEEMKEEMRHDEVLKAWAK 772

RESULT 12.

A91140

aerobic respiration sensor-response protein [imported] - Escherichia coli (strain O157:H7)

C/Species: Escherichia coli

C/Date: 18-Feb-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C/Accession: A91140

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gaaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Rep. 8, 11-22, 2001

A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend

A/Reference number: A9629; PMID:21156231; PMID:11258796

A/Accession: A91140

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-778 <HAV>

A/Cross-references: GB:BA000007; PIDN:BAH37512.1; PID:G13363562; GSPDB:GN00154

A/Experimental source: strain O157:H7, substrain RMD 0509952

C/Genetics:

A/Gene: Ecg4089

C/Superfamily: aerobic respiration control sensor protein arcb, response regulator homol

Query Match 29.5%; Score 56; DB 2; Length 778;

Best Local Similarity 32.0%; Pred. No. 25;

Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 12 VRRVRRVRRVRRVRRVRRVRR 36
DB 750 VGEWIEEMKEEMRHDEVLKAWAK 774

RESULT 13

D85985

aerobic respiration sensor-response protein [imported] - Escherichia coli (strain O157:H7)

C/Species: Escherichia coli

C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001

C/Accession: D85985

R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, J.; Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001

A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A/Reference number: A85480; PMID:21074935; PMID:11206551

A/Accession: D85985

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-778 <STO>

A/Cross-references: GB:AE005174; NID:G12517831; PIDN:AA658344.1; GSPDB:GN00145; UMGP:Z45

C/Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

A/Gene: arcb

C/Superfamily: aerobic respiration control sensor protein arcb, response regulator homol

Query Match 29.5%; Score 56; DB 2; Length 778;

Best Local Similarity 32.0%; Pred. No. 25;

Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 12 VRRVRRVRRVRRVRRVRRVRR 36
DB 750 VGEWIEEMKEEMRHDEVLKAWAK 774

RESULT 14

F90177

hypothetical protein secE [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 17-May-2002

C/Accession: F90177

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-ung, I.; Ueffing, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, I. arett, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.

submitted to Genbank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: F90177

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-62 <KIR>

A/Cross-references: GB:AE006641; NID:G13813489; PIDN:AAK40677.1; GSPDB:GN00155

C/Genetics:

A/Gene: secE

C/Superfamily: yeast SSI protein

RESULT 15

A27741

rRNA (adenine-N6-)-methyltransferase (EC 2.1.1.48) - Streptomyces fradiae

C/Species: Streptomyces fradiae

C/Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 05-May-2000

C/Accession: A27741

R/Kamimura, S.; Weisblum, B. J. Bacteriol. 170, 1800-1811, 1988

A/Title: Translational attenuation control of ermSF, an inducible resistance determinant

A/Reference number: A27741; PMID:88169508; PMID:3127381

A/Accession: A27741

A/Molecule type: DNA

A/Residues: 1-319 <KAM>

A/Cross-references: GB:M19269; NID:G153251; PIDN:AAA26742.1; PID:G153253

C/Genetics:

A/Gene: erm SF

C/Superfamily: rRNA (adenine-N6-)-methyltransferase

C/Keywords: methyltransferase; S-adenosylmethionine

Query Match 28.9%; Score 55; DB 2; Length 319;

Best Local Similarity 45.2%; Pred. No. 15;

Matches 14; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 2 RRVRRVRRVRRVRRVRRVRRVRR 32
DB 208 RHEWRLGVRVRERPRVDSGLIRER 238

Search completed: June 9, 2003, 12:03:17
Job time: 16.9362 secs

RESULT 7

T17699
 arginine-rich protein a209R - Chlorella virus PBCV-1
 C/Species: Chlorella virus PBCV-1
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T17699
 R/Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL Data Library, May 1999
 A/Reference number: Z18806
 A/Accession: T17699
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-114 <GRA>
 A/Cross-references: EMBL:U42580; NID:G4028896; PIDN:AACG6577.1
 A/Experimental source: specific host Chlorella strain NC64A
 C/Genetics:
 A/Note: a209R

Query Match

Best Local Similarity 30.8%; Score 58.5; DB 2; Length 114;
 Pred. No. 2.3;
 Matches 17; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

2 RRVRRVRRVRRVRRVRRVRRVRRVRRVRR
 52 RNVLR-RRLRRVLRRLRRRLRRVLRRLRR 85

RESULT 8

H72708
 probable ribosomal protein L31 APE1087 - Aeropyrum pernix (strain K1)
 C/Species: Aeropyrum pernix
 C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 02-Aug-2002
 C/Accession: H72708
 R/Kawabuchi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hattawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
 A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
 A/Reference number: A72450; MUID:99310339; PMID:10382966
 A/Accession: H72708
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-105 <KAW>
 A/Cross-references: DDBJ:AB000060; NID:G5104188; PIDN:BAAB0072.1; PID:dl043858; PID:G510
 A/Experimental source: strain K1
 C/Genetics:
 A/Note: APE1087
 C/Superfamily: rat ribosomal protein L31

Query Match

Best Local Similarity 30.5%; Score 58; DB 2; Length 105;
 Pred. No. 2.4;
 Matches 17; Conservative 3; Mismatches 11; Indels 2; Gaps 2;

5 WRRVRRVRRVRRVRRVRRVRRVRRVRR 36
 8 WYVVNLRRVRRVRRVRRVRRVRRVRR 39

RESULT 9

A26512
 carb protein - Streptomyces sp.
 C/Species: Streptomyces sp.
 C/Date: 11-Mar-1988 #sequence_revision 11-Mar-1988 #text_change 18-Jun-1999
 C/Accession: A26512
 R/Epp, J.K.; Buggett, S.G.; Schoner, B.E.
 Gene 53, 73-83, 1987
 A/Title: Cloning and nucleotide sequence of a carbomycin-resistance gene from Streptomyces
 A/Reference number: A26512; MUID:87248111; PMID:3036668
 A/Accession: A26512
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-299 <EPP>
 A/Cross-references: GB:U16503; NID:G153199; PIDN:AACG2026.1; PID:G153200
 C/Superfamily: rRNA (adenine-N6)-methyltransferase

C/Keywords: antibiotic resistance

Query Match 30.0%; Score 57; DB 2; Length 299;
 Best Local Similarity 46.4%; Pred. No. 8.1;
 Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

5 WRRVRRVRRVRRVRRVRRVRRVRRVRR 32
 198 WRWGERISRRWFRPVPVDSAVLRERR 225

RESULT 10

T08479
 plasmid maintenance protein kleg [imported] - Enterobacter aerogenes
 C/Species: Enterobacter aerogenes
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C/Accession: T08479
 R/Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
 Plasmid 36, 95-111, 1996
 A/Title: Conservation of the genetic switch between replication and transfer genes of I
 A/Reference number: Z16434; MUID:97118926; PMID:8954881
 A/Accession: T08479
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-89 <THO>
 A/Cross-references: EMBL:U67194; NID:G1572520; PIDN:AACG4423.1; PID:G1572528
 C/Genetics:
 A/Note: kleg

Query Match 29.5%; Score 56; DB 2; Length 89;
 Best Local Similarity 38.2%; Pred. No. 3.6;
 Matches 13; Conservative 4; Mismatches 11; Indels 6; Gaps 1;

2 RRVRRVRRVRRVRRVRRVRRVRRVRRVRR 35
 37 RRRWRPAGRVLRPGRRTGQWLR-----ARWIR 64

RESULT 11

RGECAR
 aerobic respiration control sensor protein arcB (EC 2.7.3.-) - Escherichia coli (strain
 C/Species: Escherichia coli
 C/Date: 31-Dec-1991 #sequence_revision 17-Oct-1997 #text_change 01-Mar-2002
 C/Accession: D65112; J00295; S11794
 R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A/Title: The complete genome sequence of Escherichia coli K-12.
 A/Reference number: A64720; MUID:97426617; PMID:9278503
 A/Accession: D65112
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-776 <BLAT>
 A/Cross-references: GB:AE000400; GB:U00096; NID:G2367203; PIDN:AAC76242.1; PID:G1789603
 A/Experimental source: strain K-12, substrain MG1655
 R/Iuchi, S.; Matsuda, Z.; Fujiwara, T.; Itoh, E.C.C.
 Mol. Microbiol. 4, 715-727, 1990
 A/Title: The arcB gene of Escherichia coli encodes a sensor-regulator protein for anaer
 A/Reference number: J00295; MUID:90355832; PMID:2201868
 A/Accession: J00295
 A/Molecule type: DNA
 A/Residues: 1-468; TG, 469-776 <IUC>
 A/Cross-references: EMBL:X53315; NID:G40950; PIDN:CAA37397.1; PID:G40951
 C/Genetics:
 A/Note: arcB
 A/Map position: 69.5 min
 C/Superfamily: aerobic respiration control sensor protein arcB, response regulator homo
 C/Keywords: autophosphorylation; phosphotransferase; phosphoprotein; phosphotransferase;
 F:23-50/Domain: transmembrane #status predicted <TM>
 F:59-77/Domain: transmembrane #status predicted <TM>
 F:78-776/Domain: intracellular #status predicted <INT>
 F:526-637/Domain: response regulator homology <RRH>
 F:292/Binding site: phosphate (His) (covalent) #status predicted

LIBRARY: MMLR1DT01
CLONE: 472480
US-09-016-000-4

Query Match 33.8% Score 44; DB 2; Length 556;
Best Local Similarity 41.7%; Pred. No. 2.6e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 RWVRRVRRVRR 13
Db 534 KMCKRIQEVWRQ 545

Search completed: June 9, 2003, 12:05:09
Job time : 9.17021 secs

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,891A
FILING DATE: 25-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,400
FILING DATE: 10-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-293625
FILING DATE: 14-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: KP-7501
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 525 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-348-891A-2

Query Match 34.2%; Score 44.5; DB 1; Length 525;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 10; Conservative 7; Mismatches 3; Indels 5; Gaps 2;

QY 1 RRVRRRVRRVRRV---RVRRRV 22
DB 194 RRMVKTQQ--RRVGEFRLERKWL 216

RESULT 14
US-08-905-817-2
Sequence 2, Application US/08905817
Patent No. 5824777
GENERAL INFORMATION:
APPLICANT: SASAKI, Keiko
APPLICANT: MORI, Takayuki
APPLICANT: MAKINO, Satoshi
TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE
TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,817
FILING DATE: 04-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,891
FILING DATE: 25-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,400
FILING DATE: 10-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-293625
FILING DATE: 14-OCT-1991

ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: KP-7501A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 525 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-905-817-2

Query Match 34.2%; Score 44.5; DB 2; Length 525;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 10; Conservative 7; Mismatches 3; Indels 5; Gaps 2;

QY 1 RRVRRRVRRVRRV---RVRRRV 22
DB 194 RRMVKTQQ--RRVGEFRLERKWL 216

RESULT 15
US-09-016-000-4
Sequence 4, Application US/09016000
Patent No. 5962232
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Bandman, Olga
APPLICANT: Akherblom, Ingrid E.
APPLICANT: Shah, Purni
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
TITLE OF INVENTION: PROTEIN KINASE MOLECULES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,000
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0465 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:

TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE: N/A
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-17

Query Match 35.4%; Score 46; DB 2; Length 38;
Best Local Similarity 55.0%; Pred. No. 14;
Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 RRMVRRVRRVRRVRRVRR 20
Db 3 RRAARRRARRARRARRARR 22

RESULT 11
US-08-436-703B-5
; Sequence 5, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; STREET: Suite 1525
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44MB, 3.5"
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6;
; SOFTWARE: ASCII (DOS) Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7MK-060548-00233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-965-1976
; TELEFAX: 313-965-1951
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE: N/A
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A

US-08-436-703B-5

Query Match 35.4%; Score 46; DB 2; Length 39;
Best Local Similarity 55.0%; Pred. No. 14;
Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 RRMVRRVRRVRRVRRVRR 20
Db 3 RRAARRRARRARRARRARR 22

RESULT 12

US-08-849-486-9
; Sequence 9, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
NUMBER OF SEQUENCES: 10
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,486
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95 11714
FILING DATE: 05-OCT-1995
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-849-486-9

Query Match 34.6%; Score 45; DB 3; Length 16;
Best Local Similarity 69.2%; Pred. No. 8.3;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RRMVRRVRRVRR 13
Db 1 RRMVRRVRRVRR 13

RESULT 13

US-08-348-891A-2
; Sequence 2, Application US/08348891A
; Patent No. 5654136
; GENERAL INFORMATION:

APPLICANT: SASAKI, Keiko
APPLICANT: MORI, Takayuki
APPLICANT: MAKINO, Satoshi
TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE,
CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
TITLE OF INVENTION: ITS ABSOLUTE IDENTIFICATION
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

APPLICATION NUMBER: PCT/US95/00062
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Yates, Michael E.; Sweeney, Patricia A.;
NAME: Roth, Michael J.; & Simon, Soma G.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 234R2-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00062-9

Query Match 37.7%; Score 49; DB 5; Length 31;
Best Local Similarity 45.0%; Pred. No. 5.1;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 RRVRRVRRVRRVRRVRR 24
DB 1 RRVRRVRRVRRVRRVRR 20

RESULT 8
US-09-413-814-78
Sequence 78, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Geesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Bayer, Stefan
APPLICANT: Bioecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
TITLE OF INVENTION: heteropolypeptide compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
EARLIER FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 78
LENGTH: 882
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-78

Query Match 37.3%; Score 48.5; DB 4; Length 882;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 15; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 RRVRRVRRVRRVRRVRRVRR 24
DB 601 RRVRRVRRVRRVRRVRRVRR 625

RESULT 9

US-09-092-315-8
Sequence 8, Application US/09092315
Patent No. 6399337
GENERAL INFORMATION:
APPLICANT: Taylor, Diane E.
APPLICANT: Ge, Zhongming
TITLE OF INVENTION: ALPHA-1, 3-FUCOSYLTRANSFERASE
FILE REFERENCE: 07254/049001
CURRENT APPLICATION NUMBER: US/09/092,315
EARLIER FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: US 60/048,857
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 454
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-092-315-8

Query Match 35.8%; Score 46.5; DB 4; Length 454;
Best Local Similarity 35.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 10; Mismatches 0; Indels 3; Gaps 1;

QY 8 RRVRRVRRVRRVRRVRR 24
DB 435 RRVRRVRRVRRVRRVRR 454

RESULT 10
US-08-436-703B-17
Sequence 17, Application US/08436703B
Patent No. 5919761
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
APPLICANT: Stanley, James C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR
TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
TITLE OF INVENTION: WEIGHT HEPARIN
TITLE OF INVENTION: ANTICONGULANT REVERSAL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSER: Benita J. Rohm, Esq.
STREET: 6601 Woodward Avenue
CITY: Suite 1525
CITY: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6;
SOFTWARE: ASCII (DOS) text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7W1-060548-00233
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids

CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 2618
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-28

Query Match 38.5%; Score 50; DB 4; Length 2618;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 12; Conservative 0; Mismatches 6; Indels 6; Gaps 1;

QY 4 RRRRRRVVV-----RVRRRM 21
DB 48 VAFRRARVVAADQLRAVFRM 71

RESULT 5
US-08-179-632-9
Sequence 9, Application US/08179632
Patent No. 5607914
GENERAL INFORMATION:
APPLICANT: Rao, A. Gururaj; Zhong, Lingxiu
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/Microsoft Windows
SOFTWARE: Microsoft Windows No. 5607914epad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/179,632
FILING DATE: 07-JAN-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/079,512
FILING DATE: 06/18/93
ATTORNEY/AGENT INFORMATION:
NAME: Roch, Michael J.
REGISTRATION NUMBER: 29,342
REFERENCE/DOCKET NUMBER: 0233 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 245-3594
TELEFAX: (515) 245-3634
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-179-632-9

Query Match 37.7%; Score 49; DB 1; Length 31;
Best Local Similarity 45.0%; Pred. No. 5.1;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 RRVRRRVVVRRVRRVR 24
DB 1 RRIYRAIRHPRIRIGWLR 20

RESULT 6
US-08-440-174A-9
Sequence 9, Application US/08440174A
Patent No. 5717061
GENERAL INFORMATION:
APPLICANT: Rao, Gururaj A.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: 7100 N.W. 62nd Avenue
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,174A
FILING DATE: 12-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/079,512
FILING DATE: 18-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bobrowicz, Donna
REGISTRATION NUMBER: 32,196
REFERENCE/DOCKET NUMBER: 0234R2D-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4896
TELEFAX: (515) 334-6883

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-174A-9

Query Match 37.7%; Score 49; DB 1; Length 31;
Best Local Similarity 45.0%; Pred. No. 5.1;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 RRVRRRVVVRRVRRVR 24
DB 1 RRIYRAIRHPRIRIGWLR 20

RESULT 7
PCT-US95-00062-9
Sequence 9, Application PC/TUS9500062
GENERAL INFORMATION:
APPLICANT: Pioneer Hi-Bred International, Inc.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/Microsoft Windows
SOFTWARE: Microsoft Windows Notepad
CURRENT APPLICATION DATA:

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;; ADDRESSEE: P.C.
;; STREET: 1755 S. Jefferson Davis Highway, Suite 400
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Releasee #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/286,819A
;; FILING DATE: 05-AUG-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/174,682
;; FILING DATE: 28-DEC-1993
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/917,146
;; FILING DATE: 10-AUG-1992
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/FR/91/00855
;; FILING DATE: 29-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 9013579
;; FILING DATE: 31-OCT-1990
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Oblon, No. 5871910man F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-060-0 PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2254 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-286-819A-28
;;
Query Match 39.2%; Score 51; DB 2; Length 2254;
Best Local Similarity 45.5%; Pred. No. 1.4e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 2 RMVRRVRRVRRVRRVRRVRR 23
DB 1627 RAVRSVRYWETVPRLARWTK 1648
;;
RESULT 3
;; US-08-980-357-28
;; Sequence 28, Application US/08980357
;; Patent No. 6013508
;; GENERAL INFORMATION:
;; APPLICANT: ARTHUR, MICHEL,
;; APPLICANT: DUKTA-WALEN, SYLVIE
;; APPLICANT: MOLINIS, CATHERINE
;; APPLICANT: COURVALIN, PATRICE
;; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
;; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEIDS, IN PARTICULAR
;; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
;; TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
;; NUMBER OF SEQUENCES: 54
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

;; ADDRESSEE: P.C.
;; STREET: 1755 S. Jefferson Davis Highway, Suite 400
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Releasee #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/980,357
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/286,819
;; FILING DATE: 05-AUG-1994
;; APPLICATION NUMBER: US 08/174,682
;; FILING DATE: 28-DEC-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/917,146
;; FILING DATE: 10-AUG-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/FR/91/00855
;; FILING DATE: 29-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 9013579
;; FILING DATE: 31-OCT-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Oblon, No. 6013508man F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-060-0 PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2254 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-980-357-28
;;
Query Match 39.2%; Score 51; DB 3; Length 2254;
Best Local Similarity 45.5%; Pred. No. 1.4e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 2 RMVRRVRRVRRVRRVRRVRR 23
DB 1627 RAVRSVRYWETVPRLARWTK 1648
;;
RESULT 4
;; US-09-413-814-28
;; Sequence 28, Application US/09413814
;; Patent No. 6225064
;; GENERAL INFORMATION:
;; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
;; APPLICANT: Bristol-Myers Squibb, Co.
;; APPLICANT: Beyer, Stefan
;; APPLICANT: Bioecker, Helmut
;; APPLICANT: Brandt, Petra
;; APPLICANT: Cino, Paul M
;; APPLICANT: Dougherty, Brian A
;; APPLICANT: Goldberg, Steven L
;; APPLICANT: Hofle, Gerhard
;; APPLICANT: Meiller, Joachim
;; APPLICANT: Reichenbach, Hans
;; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
;; TITLE OF INVENTION: heteropolyketide compounds
;; FILE REFERENCE: PCT/US 99/23535

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:50:31 ; Search time 8.17021 Seconds
(without alignments)
86,430 Million cell updates/sec

Title: US-10-079-075-10

Perfect score: 130
Sequence: 1 RRVRRRRRRRRRRRRRRRRRRRR

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A-COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B-COMB.pep:*
5: /cgn2_6/prodata/1/iaa/pctus-COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfillseq.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	40.0	770	4	US-08-245-248B-31
2	51	39.2	2254	2	US-08-286-819A-28
3	51	39.2	2254	3	US-08-980-357-28
4	50	38.5	2618	4	US-09-413-814-28
5	49	37.7	31	1	US-08-179-632-9
6	49	37.7	31	1	US-08-440-174A-9
7	49	37.7	31	5	PCT-US95-00062-9
8	48.5	37.3	882	4	US-09-413-814-78
9	46.5	35.8	454	4	US-09-092-315-8
10	46	35.4	38	2	US-08-436-703B-17
11	46	35.4	39	2	US-08-436-703B-5
12	45	34.6	16	3	US-08-849-486-9
13	44.5	34.2	525	1	US-08-348-891A-2
14	44.5	34.2	525	2	US-08-305-817-2
15	44	33.8	556	2	US-09-016-000-4
16	44	33.8	1088	2	US-08-742-026-2
17	44	33.8	1088	2	US-08-742-026-23
18	43	33.1	27	2	US-08-505-486-65
19	43	33.1	27	2	US-08-801-028-65
20	43	33.1	27	3	US-09-340-154-65
21	43	33.1	27	4	US-09-482-611B-65
22	43	33.1	27	5	PCT-US95-09338-65
23	43	33.1	27	5	PCT-US95-09339-65
24	42	32.3	114	4	US-09-055-113-3
25	42	32.3	777	4	US-09-257-770-7
26	42	32.3	2205	1	US-08-093-453B-2
27	41.5	31.9	48	2	US-08-637-759B-253

28	41.5	31.9	48	3	US-08-871-355A-253	Sequence 253, App
29	41.5	31.9	48	4	US-09-201-945-253	Sequence 253, App
30	41.5	31.9	879	1	US-08-220-151-2	Sequence 2, Appl
31	41.5	31.9	879	1	US-08-220-151-3	Sequence 2, Appl
32	41.5	31.9	879	1	US-08-413-118-2	Sequence 2, Appl
33	41.5	31.9	879	1	US-08-413-118-3	Sequence 3, Appl
34	41.5	31.9	879	1	US-08-413-118-106	Sequence 106, App
35	41.5	31.9	879	3	US-08-473-446-2	Sequence 2, Appl
36	41.5	31.9	879	3	US-08-473-446-3	Sequence 106, App
37	41.5	31.9	879	3	US-08-473-446-106	Sequence 106, App
38	41	31.5	21	1	US-08-786-748A-50	Sequence 50, Appl
39	41	31.5	21	1	US-08-932-682-50	Sequence 50, Appl
40	41	31.5	21	2	US-08-932-682-53	Sequence 50, Appl
41	41	31.5	21	2	US-08-932-682-53	Sequence 50, Appl
42	41	31.5	23	2	US-08-505-486-63	Sequence 63, Appl
43	41	31.5	23	3	US-08-801-028-63	Sequence 63, Appl
44	41	31.5	23	3	US-09-340-154-63	Sequence 63, Appl
45	41	31.5	23	4	US-09-482-611B-63	Sequence 63, Appl

ALIGNMENTS

```
RESULT 1
US-09-245-248B-31
; Sequence 31, Application US/09245248B
; Patent No. 6395472
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Leary, Thomas
; APPLICANT: Erker, James
; APPLICANT: Chalmers, Michelle
; APPLICANT: Simons, John
; APPLICANT: Birkenmeyer, Larry
; APPLICANT: Muerhoff, Scott
; APPLICANT: Pilot-Matias, Tami
; APPLICANT: Desai, Suresh
; APPLICANT: Mushahwar, Isa
; TITLE OF INVENTION: METHODS OF UTILIZING THE TT VIRUS
; FILE REFERENCE: 6461 US.OI
; CURRENT APPLICATION NUMBER: US/09/245,248B
; CURRENT FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-245-248B-31

Query Match          40.0%; Score 52; DB 4; Length 770;
Best local similarity 59.1%; Pred. No. 40;
Matches 13; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY      3 WVRVRVRVRVRVRVRVRVRVR 24
          ||| ||| ||| ||| |||
Db       5 WVRVRVRVRVRVRVR--RPVRRR 24

RESULT 2
US-08-286-819A-28
; Sequence 28, Application US/08286819A
; Patent No. 5871910
; GENERAL INFORMATION:
; APPLICANT: ARTHUR, MICHEL
; APPLICANT: DURKA-WALEN, SYLVIE
; APPLICANT: MOLINAS, CATHERINE
; APPLICANT: COURVALIN, PATRICE
; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
; TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
; NUMBER OF SEQUENCES: 54
```



```
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16830
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-16830
```

```
Query Match          40.0%; Score 52; DB 7; Length 142;
Best Local Similarity 44.8%; Pred. No. 2e+02; 7; Indels 6; Gaps 2;
Matches 13; Conservative 3; Mismatches 7; Indels 6; Gaps 2;
```

```
QY      1 RRVRRRVRRV-RRVR-----VVRRVVR 23
          |||:| | | | | | | | | | | | | | | |
DB      46 RRVRRSQRTYCRRMWRMSERTSCRRWR 74
```

```
RESULT 15
US-10-366-683-20533
; Sequence 20533, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolling, Jock
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20533
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-20533
```

```
Query Match          40.0%; Score 52; DB 6; Length 526;
Best Local Similarity 69.2%; Pred. No. 5.8e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 RRVRRRVRRVR 13
          ||| | | | | |
DB      1 RRVTRPARRRMR 13
```

Search completed: June 9, 2003, 12:31:38
Job time : 32.0638 secs

PRIOR FILING DATE: 1999-02-18
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 16956
LENGTH: 387
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-366-683-16956

Query Match 46.5%; Score 60.5; DB 6; Length 387;
Best Local Similarity 38.6%; Pred. No. 58;
Matches 17; Conservative 1; Mismatches 5; Indels 21; Gaps 2;

QY 1 RRVRRVR--VW-----RRVRRVRWR 23
DB 309 RMLPRGRSCPWTGCGGSGPAVAVRAGRTTAVRRVRWR 352

RESULT 7
US-10-419-128-16956
Sequence 16956, Application US/10419128
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/10/419,128
CURRENT FILING DATE: 2003-04-21
PRIOR APPLICATION NUMBER: US/09/252,991
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 16956
LENGTH: 387
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-419-128-16956

Query Match 46.5%; Score 60.5; DB 6; Length 387;
Best Local Similarity 38.6%; Pred. No. 58;
Matches 17; Conservative 1; Mismatches 5; Indels 21; Gaps 2;

QY 1 RRVRRVR--VW-----RRVRRVRWR 23
DB 309 RMLPRGRSCPWTGCGGSGPAVAVRAGRTTAVRRVRWR 352

RESULT 8
US-10-425-114-53360
Sequence 53360, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 53360
LENGTH: 107
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700155612_FLI.pep
US-10-425-114-53360

Query Match 44.2%; Score 57.5; DB 6; Length 107;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 15; Conservative 1; Mismatches 5; Indels 9; Gaps 2;
QY 2 RRVRRVRWRVVR-----VVRWR 24
DB 45 RVRRRRDVRRRRRAVGLARRVVR 72

RESULT 9
US-10-425-114-56955
Sequence 56955, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 56955
LENGTH: 107
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3245-483-D5_FLI.pep
US-10-425-114-56955

Query Match 44.2%; Score 57.5; DB 6; Length 107;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 15; Conservative 1; Mismatches 5; Indels 9; Gaps 2;

QY 2 RRVRRVRWRVVR-----VVRWR 24
DB 45 RVRRRRDVRRRRRAVGLARRVVR 72

RESULT 10
US-10-264-237-2376
Sequence 2376, Application US/10264237
GENERAL INFORMATION:
APPLICANT: Bires et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P4131P1
CURRENT APPLICATION NUMBER: US/10/264,237
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/16450
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,515
PRIOR FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 2876
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2376
LENGTH: 136
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (90)_FEATURE
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-2376

Query Match 43.8%; Score 57; DB 6; Length 136;
Best Local Similarity 40.9%; Pred. No. 59;
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 3 VRRVRRVRRVRRVRRVRRVRR 24
US-10-264-237-2376


```
; SEQ ID NO 58367
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3591-091-B12_F11.pcp
US-10-425-114-58367
```

```
Query Match          47.7%; Score 62; DB 6; Length 129;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 12; Conservative 1; Mismatches 1; Indels 6; Gaps 1;
```

```
Qy          2 RWRVRVRVRVRVRVRVRVR 21
              |||:|:|:|:|:|:|:|
Db          12 RWRRLRRRRR-----RRW 25
```

```
RESULT 3
US-10-017-161-2262
; Sequence 2262, Application US/10017161
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789.
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2262
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (5)_RES
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (8)_RES
; OTHER INFORMATION: Variable amino acid
; NAME/KEY: MOD_RES
; LOCATION: (17)
; OTHER INFORMATION: Variable amino acid
US-10-017-161-2262
```

```
Query Match          47.7%; Score 62; DB 6; Length 387;
Best Local Similarity 63.6%; Pred. No. 41;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy          1 RWRVRVRVRVRVRVRVRVR 22
              |||:|:|:|:|:|:|:|
Db          315 RWRVRAGSVVRWVGAVSRWV 336
```

```
RESULT 4
US-10-219-051B-8227
; Sequence 8227, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
```

```
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 8227
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: SWISS-Prot / AAA68695
; DATABASE ENTRY DATE: 1998-11-01
US-10-219-051B-8227
```

```
Query Match          47.7%; Score 62; DB 6; Length 396;
Best Local Similarity 47.8%; Pred. No. 41;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
```

```
Qy          2 RWRVRVRVRVRVRVRVRVR 24
              |||:|:|:|:|:|:|:|
Db          107 RWRKREHWVRREVFYRLRRWADR 129
```

```
RESULT 5
US-10-219-051B-8229
; Sequence 8229, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 8229
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Refseq / NP_056008
; DATABASE ENTRY DATE: 2002-10-31
US-10-219-051B-8229
```

```
Query Match          47.7%; Score 62; DB 6; Length 396;
Best Local Similarity 47.8%; Pred. No. 41;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
```

```
Qy          2 RWRVRVRVRVRVRVRVRVR 24
              |||:|:~|:|:~|:~|:~|
Db          107 RWRKREHWVRREVFYRLRRWADR 129
```

```
RESULT 6
US-10-366-683-16956
; Sequence 16956, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubinfeld, Marc J.
; APPLICANT: Nolleng, Jork
; APPLICANT: Deloughery, Craig
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PAT93-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:57:02 ; Search time 31.0638 Seconds
(without alignments)
173.908 Million cell updates/sec

Title: US-10-079-075-10

Perfect score: 130

Sequence: 1 RRVRRRRRRRRRRRRRRRRRRRRR 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1046584 seqs, 225093350 residues

Total number of hits satisfying chosen parameters: 1046584

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA.New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73.5	56.5	245	US-10-425-114-70663	Sequence 70663, A
2	62	47.7	129	US-10-425-114-58367	Sequence 58367, A
3	62	47.7	387	US-10-017-161-2262	Sequence 2262, AP
4	62	47.7	396	US-10-219-051B-8227	Sequence 8227, AP
5	62	47.7	396	US-10-219-051B-8229	Sequence 8229, AP
6	60.5	46.5	387	US-10-366-683-16956	Sequence 16956, A
7	60.5	46.5	387	US-10-419-128-16956	Sequence 16956, A
8	57.5	44.2	107	US-10-425-114-53360	Sequence 53360, A
9	57.5	44.2	107	US-10-425-114-56955	Sequence 56955, A
10	57	43.1	139	US-10-264-237-2376	Sequence 2376, AP
11	56	43.1	139	US-10-425-114-52367	Sequence 52367, A
12	53.5	41.2	360	US-10-425-114-66861	Sequence 66861, A
13	52.5	40.4	360	US-10-282-122A-49117	Sequence 49117, A
14	52	40.0	142	US-60-452-680-16830	Sequence 16830, A
15	52	40.0	526	US-10-366-683-20533	Sequence 20533, A
16	52	40.0	526	US-10-419-128-20533	Sequence 20533, A
17	51.5	39.6	181	US-10-425-114-60359	Sequence 60359, A
18	51.5	39.6	181	US-10-425-114-48236	Sequence 48236, A
19	51.5	39.6	199	US-10-425-114-53496	Sequence 53496, A
20	51	39.2	149	US-10-424-599-154312	Sequence 154312, A
21	51	39.2	195	US-10-425-114-70810	Sequence 70810, A
22	51	39.2	195	US-10-425-114-68513	Sequence 68513, A
23	51	39.2	209	US-10-425-114-68542	Sequence 68542, A
24	51	39.2	245	US-10-282-122A-53812	Sequence 53812, A
25	51	39.2	252	US-10-425-114-71061	Sequence 71061, A
26	51	39.2	342	US-10-425-114-68594	Sequence 68594, A

27	51	39.2	570	US-10-369-493-5081	Sequence 5081, AP
28	51	39.2	957	US-10-282-122A-64361	Sequence 64361, A
29	50.5	38.8	77	US-10-424-599-179654	Sequence 179654, A
30	50.5	38.8	91	US-10-424-599-265356	Sequence 265356, A
31	50.5	38.8	549	US-10-438-246-17783	Sequence 17783, A
32	50.5	38.8	585	US-10-438-246-17731	Sequence 17731, A
33	50	38.5	28	US-09-874-644A-17	Sequence 17, Appl
34	50	38.5	106	US-10-289-762-253	Sequence 253, Appl
35	50	38.5	150	US-09-675-784A-8983	Sequence 8983, AP
36	50	38.5	166	US-10-417-086-7345	Sequence 7345, AP
37	50	38.5	288	US-10-369-493-17910	Sequence 17910, A
38	50	38.5	334	US-10-378-029-62	Sequence 62, Appl
39	50	38.5	473	US-10-446-203-8968	Sequence 8968, AP
40	50	38.5	692	US-10-156-761-12598	Sequence 12598, A
41	49.5	38.1	76	PCT-US02-32727-13379	Sequence 13379, A
42	49.5	38.1	76	US-09-978-825-13379	Sequence 13379, A
43	49.5	38.1	76	US-10-057-498-13379	Sequence 13379, A
44	49.5	38.1	200	US-10-366-683-28054	Sequence 28054, A
45	49.5	38.1	200	US-10-419-128-28054	Sequence 28054, A

ALIGNMENTS

```
RESULT 1
US-10-425-114-70663
; Sequence 70663, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70663
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73064A07_Flt.pep
US-10-425-114-70663

Query Match      56.5%; Score 73.5; DB 6; Length 245;
Best Local Similarity 56.7%; Pred. No. 1.8;
Matches 17; Conservative 1; Mismatches 5; Indels 7; Gaps 1;

CY      2 RRVRRRRRRR-----RVVRVRRVRR 24
      |||||:|||||
DB      12 RRVRRRRRRRRRRRRRRRRRRRRRR 41

RESULT 2
US-10-425-114-58367
; Sequence 58367, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
```

```
; Sequence 12, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
PCT-US02-04812-12
```

```
Query Match          100.0%; Score 130; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30
```

```
RESULT 13
US-09-785-058-12
; Sequence 12, Application US/09785058
```

```
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-058-12
```

```
Query Match          100.0%; Score 130; DB 21; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30
```

```
RESULT 14
```

```
US-09-785-059-12
; Sequence 12, Application US/09785059
```

```
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30
```

```
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-059-12
```

```
Query Match          100.0%; Score 130; DB 21; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30
```

```
RESULT 15
```

```
US-10-079-075-12
; Sequence 12, Application US/10079075
```

```
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-10-079-075-12
```

```
Query Match          100.0%; Score 130; DB 24; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30
```

```
Search completed: June 9, 2003, 12:25:29
Job time : 103.915 secs
```

```
PCT-US02-04812-11
; Sequence 11, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-11

Query Match          100.0%; Score 130; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 13 RRVRRRVRRRVRRRVRRRVRR 36

RESULT 8
US-09-785-058-11
; Sequence 11, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-11

Query Match          100.0%; Score 130; DB 21; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 13 RRVRRRVRRRVRRRVRRRVRR 36

RESULT 9
US-09-785-059-11
; Sequence 11, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
```

```
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-11

Query Match          100.0%; Score 130; DB 21; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 13 RRVRRRVRRRVRRRVRRRVRR 36

RESULT 10
US-10-079-075-11
; Sequence 11, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-11

Query Match          100.0%; Score 130; DB 24; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 13 RRVRRRVRRRVRRRVRRRVRR 36

RESULT 11
PCT-US02-04432-12
; Sequence 12, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
PCT-US02-04432-12

Query Match          100.0%; Score 130; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 7 RRVRRRVRRRVRRRVRRRVRR 30

RESULT 12
PCT-US02-04812-12
```

```
RESULT 2
PCT-US02-04812-10
; Sequence 10, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-10

Query Match
Best Local Similarity 100.0%; Score 130; DB 1; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVVRRVRRVRRVRRVRRVRR 24
Db 1 RRVVRRVRRVRRVRRVRRVRR 24

RESULT 3
US-09-785-058-10
; Sequence 10, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-10

Query Match
Best Local Similarity 100.0%; Score 130; DB 21; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVVRRVRRVRRVRRVRRVRR 24
Db 1 RRVVRRVRRVRRVRRVRRVRR 24

RESULT 4
US-09-785-059-10
; Sequence 10, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-10

Query Match
Best Local Similarity 100.0%; Score 130; DB 21; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVVRRVRRVRRVRRVRRVRR 24
Db 1 RRVVRRVRRVRRVRRVRRVRR 24

RESULT 5
US-10-079-075-10
; Sequence 10, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-10

Query Match
Best Local Similarity 100.0%; Score 130; DB 24; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVVRRVRRVRRVRRVRRVRR 24
Db 1 RRVVRRVRRVRRVRRVRRVRR 24
```

```
RESULT 6
PCT-US02-04432-11
; Sequence 11, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04432-11

Query Match
Best Local Similarity 100.0%; Score 130; DB 1; Length 36;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVVRRVRRVRRVRRVRRVRR 24
Db 13 RRVVRRVRRVRRVRRVRRVRR 36

RESULT 7
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:55:47 ; Search time 103.915 Seconds
(without alignments)
148.906 Million cell updates/sec

Title: US-10-079-075-10
Page: 120

Perfect score: 130
Sequence: 1 RRVVRRVRRVRRVRRVRR 24

Scoring code: BVSUM02
Gapof 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473310 residues

Total number of hits satisfying chosen parameters: 4569144

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

Database :

27:	/cgnt_6/prodata/1/paa/US010_COMB_dep.*
26:	/cgnt_6/prodata/1/paa/US102_COMB_dep.*
25:	/cgnt_6/prodata/1/paa/US101_COMB_dep.*
24:	/cgnt_6/prodata/1/paa/US100_COMB_dep.*
23:	/cgnt_6/prodata/1/paa/US099_COMB_dep.*
22:	/cgnt_6/prodata/1/paa/US098_COMB_dep.*
21:	/cgnt_6/prodata/1/paa/US097_COMB_dep.*
20:	/cgnt_6/prodata/1/paa/US096_COMB_dep.*
19:	/cgnt_6/prodata/1/paa/US095_COMB_dep.*
18:	/cgnt_6/prodata/1/paa/US094_COMB_dep.*
17:	/cgnt_6/prodata/1/paa/US093_COMB_dep.*
16:	/cgnt_6/prodata/1/paa/US092_COMB_dep.*
15:	/cgnt_6/prodata/1/paa/US091_COMB_dep.*
14:	/cgnt_6/prodata/1/paa/US090_COMB_dep.*
13:	/cgnt_6/prodata/1/paa/US089_COMB_dep.*
12:	/cgnt_6/prodata/1/paa/US088_COMB_dep.*
11:	/cgnt_6/prodata/1/paa/US087_COMB_dep.*
10:	/cgnt_6/prodata/1/paa/US086_COMB_dep.*
9:	/cgnt_6/prodata/1/paa/US085_COMB_dep.*
8:	/cgnt_6/prodata/1/paa/US084_COMB_dep.*
7:	/cgnt_6/prodata/1/paa/US083_COMB_dep.*
6:	/cgnt_6/prodata/1/paa/US082_COMB_dep.*
5:	/cgnt_6/prodata/1/paa/US081_COMB_dep.*
4:	/cgnt_6/prodata/1/paa/US080_COMB_dep.*
3:	/cgnt_6/prodata/1/paa/US079_COMB_dep.*
2:	/cgnt_6/prodata/1/paa/US078_COMB_dep.*
1:	/cgnt_6/prodata/1/paa/US077_COMB_dep.*
	Pending Patentes AA Main:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	130	100.0	24	1	PCT-US02-04433-10	Sequence 10, Appl
2	130	100.0	24	1	PCT-US02-04812-10	Sequence 10, Appl
3	130	100.0	24	21	US-09-785-058-10	Sequence 10, App
4	130	100.0	24	21	US-09-785-059-10	Sequence 10, App
5	130	100.0	24	24	US-10-079-075-10	Sequence 10, App
6	130	100.0	36	1	PCT-US02-04432-11	Sequence 11, Appl

7	130	100.0	36	1	PCT-US02-04812-11	Sequence 11, Appl
8	130	100.0	36	21	US-09-785-058-11	Sequence 11, Appl
9	130	100.0	36	21	US-09-785-059-11	Sequence 11, Appl
10	130	100.0	36	24	US-10-079-075-11	Sequence 11, Appl
11	130	100.0	48	1	PCT-US02-04432-12	Sequence 12, Appl
12	130	100.0	48	1	PCT-US02-04812-12	Sequence 12, Appl
13	130	100.0	48	21	US-09-785-058-12	Sequence 12, Appl
14	130	100.0	48	21	US-09-785-059-12	Sequence 12, Appl
15	130	100.0	48	24	US-10-079-075-12	Sequence 12, Appl
16	88	67.7	24	1	PCT-US02-04432-5	Sequence 5, Appl
17	88	67.7	24	1	PCT-US02-04812-5	Sequence 5, Appl
18	88	67.7	24	21	US-09-785-058-5	Sequence 5, Appl
19	88	67.7	24	21	US-09-785-059-5	Sequence 5, Appl
20	88	67.7	24	24	US-10-079-075-5	Sequence 5, Appl
21	88	67.7	36	1	PCT-US02-04432-6	Sequence 6, Appl
22	88	67.7	36	1	PCT-US02-04812-6	Sequence 6, Appl
23	88	67.7	36	21	US-09-785-058-6	Sequence 6, Appl
24	88	67.7	36	21	US-09-785-059-6	Sequence 6, Appl
25	88	67.7	36	24	US-10-079-075-6	Sequence 6, Appl
26	88	67.7	42	1	PCT-US02-04432-7	Sequence 7, Appl
27	88	67.7	42	1	PCT-US02-04812-7	Sequence 7, Appl
28	88	67.7	42	21	US-09-785-058-7	Sequence 7, Appl
29	88	67.7	42	21	US-09-785-059-7	Sequence 7, Appl
30	88	67.7	42	24	US-10-079-075-7	Sequence 7, Appl
31	88	67.7	48	1	PCT-US02-04432-8	Sequence 8, Appl
32	88	67.7	48	1	PCT-US02-04812-8	Sequence 8, Appl
33	88	67.7	48	21	US-09-785-058-8	Sequence 8, Appl
34	88	67.7	48	21	US-09-785-059-8	Sequence 8, Appl
35	88	67.7	48	24	US-10-079-075-8	Sequence 8, Appl
36	78	60.0	31	1	PCT-US02-04432-3	Sequence 3, Appl
37	78	60.0	31	1	PCT-US02-04812-3	Sequence 3, Appl
38	78	60.0	31	21	US-09-785-058-3	Sequence 3, Appl
39	78	60.0	31	21	US-09-785-059-3	Sequence 3, Appl
40	78	60.0	31	24	US-10-079-075-3	Sequence 3, Appl
41	66	50.8	133	27	US-60-614-150-1.10746	Sequence 10746,
42	66	50.8	133	27	US-60-191-637-1.10778	Sequence 10778,
43	66	50.8	133	27	US-60-191-681-8450	Sequence 8450, A
44	62	47.7	460	1	PCT-US01-08656-9621	Sequence 9621, A
45	61	46.9	12	1	PCT-US02-04432-9	Sequence 9, Appl

```

RESULT 1
PCT-US02-04432-10
; Sequence 10, Application PC/RTS02004432
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-PCT / 073396.0223
CURRENT APPLICATION NUMBER: PCT/US02/04432
CURRENT FILING DATE: 2002-02-13
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-
PCT-US02-04432-10

```

Query Match	100.0%;	Score 130;	DB 1;	Length 24;
Best Local Similarity	100.0%;	Pred. No. 3.9e-09;		
Matches	24;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0
Qy	1	RRRVRRRVRRRVRRRVRRRVRR	24	
Db	1	RRRVRRRVRRRVRRRVRRRVRR	24	

```
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Mietzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A 34001 / 072396.0222
/ CURRENT APPLICATION NUMBER: US/09/785,058
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5
/ LENGTH: 24
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-5

Query Match      67.7%; Score 88; DB 9; Length 24;
Best Local Similarity 87.5%; Pred. No. 5.4e-05;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 RRVRRRVRRVRRVRRVRRVRR 24
Db      13 RRVRRRVRRVRRVRRVRRVRR 24

RESULT 13
US-09-785-059-6
/ Sequence 6, Application US/09785059
/ Patent No. US20020169279A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Mietzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A33577 / 072396.0217
/ CURRENT APPLICATION NUMBER: US/09/785,059
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 6
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-6

Query Match      67.7%; Score 88; DB 9; Length 36;
Best Local Similarity 87.5%; Pred. No. 8.1e-05;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 RRVRRRVRRVRRVRRVRRVRR 24
Db      13 RRVRRRVRRVRRVRRVRRVRR 36

RESULT 14
US-10-079-075-6
/ Sequence 6, Application US/10079075
/ Publication No. US20020188102A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Mietzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A34001-A / 072396.0222
/ CURRENT APPLICATION NUMBER: US/10/079,075
/ CURRENT FILING DATE: 2002-02-19
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 6
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
```

```
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-6

Query Match      67.7%; Score 88; DB 9; Length 36;
Best Local Similarity 87.5%; Pred. No. 8.1e-05;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 RRVRRRVRRVRRVRRVRRVRR 24
Db      13 RRVRRRVRRVRRVRRVRRVRR 36

RESULT 15
US-09-785-058-6
/ Sequence 6, Application US/09785058
/ Publication No. US20030036627A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Mietzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A 34001 / 072396.0222
/ CURRENT APPLICATION NUMBER: US/09/785,058
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 6
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-6

Query Match      67.7%; Score 88; DB 9; Length 36;
Best Local Similarity 87.5%; Pred. No. 8.1e-05;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 RRVRRRVRRVRRVRRVRRVRR 24
Db      13 RRVRRRVRRVRRVRRVRRVRR 36

Search completed: June 9, 2003, 12:34:11
Job time : 13.766 secs
```

FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785,059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-059-12

Query Match 100.0%; Score 130; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 6,5e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30

RESULT 8
US-10-079-075-12
Sequence 12, Application US/10079075
Publication No. US2002018102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079,075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: artificial peptides derived from HIV-1
US-10-079-075-12

Query Match 100.0%; Score 130; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 6,5e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30

RESULT 9
US-09-785-058-12
Sequence 12, Application US/09785058
Publication No. US20030036627A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A 34001 / 072396.0222
CURRENT APPLICATION NUMBER: US/09/785,058
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-058-12

Query Match 100.0%; Score 130; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 6,5e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30

RESULT 10
US-09-785-059-5
Sequence 5, Application US/09785059
Patent No. US20020169279A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785,059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 24
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-5

Query Match 67.7%; Score 88; DB 9; Length 24;
Best Local Similarity 87.5%; Pred. No. 5,4e-05;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 11
US-10-079-075-5
Sequence 5, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079,075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 24
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-5

Query Match 67.7%; Score 88; DB 9; Length 24;
Best Local Similarity 87.5%; Pred. No. 5,4e-05;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 12
US-09-785-058-5
Sequence 5, Application US/09785058
Publication No. US20030036627A1


```
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 10
/ LENGTH: 24
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-10

Query Match          100.0%; Score 130; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRVRRVRRVRRVRRVRRVRR 24
Db 1 RRRVRRVRRVRRVRRVRRVRR 24

RESULT 3
US-09-785-058-10
/ Sequence 10, Application US/09785058
/ Publication No. US20030036627A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A 34001 / 072396.0222
/ CURRENT APPLICATION NUMBER: US/09/785,058
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 10
/ LENGTH: 24
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-10

Query Match          100.0%; Score 130; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRVRRVRRVRRVRRVRRVRR 24
Db 1 RRRVRRVRRVRRVRRVRRVRR 24

RESULT 4
US-09-785-059-11
/ Sequence 11, Application US/09785059
/ Patent No. US20020169279A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A33577 / 072396.0217
/ CURRENT APPLICATION NUMBER: US/09/785,059
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-11

Query Match          100.0%; Score 130; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.9e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 RRRVRRVRRVRRVRRVRRVRR 24
Db 13 RRRVRRVRRVRRVRRVRRVRR 36

RESULT 5
US-10-079-075-11
/ Sequence 11, Application US/10079075
/ Publication No. US20020189102A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A34001-A / 072396.0222
/ CURRENT APPLICATION NUMBER: US/10/079,075
/ CURRENT FILING DATE: 2002-02-19
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-11

Query Match          100.0%; Score 130; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.9e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRVRRVRRVRRVRRVRRVRR 24
Db 13 RRRVRRVRRVRRVRRVRRVRR 36

RESULT 6
US-09-785-058-11
/ Sequence 11, Application US/09785058
/ Publication No. US20030036627A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A 34001 / 072396.0222
/ CURRENT APPLICATION NUMBER: US/09/785,058
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-11

Query Match          100.0%; Score 130; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.9e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRVRRVRRVRRVRRVRRVRR 24
Db 13 RRRVRRVRRVRRVRRVRRVRR 36

RESULT 7
US-09-785-059-12
/ Sequence 12, Application US/09785059
/ Patent No. US20020169279A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 12:01:36 ; Search time 12.766 Seconds
(without alignments)
194.092 Million cell updates/sec

Title: US-10-079-075-10

Perfect score: 130

Sequence: 1 RRMVRRVRRVRRVRRVRRVRR 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:

1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB_PEP.*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB_PEP.*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB_PEP.*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB_PEP.*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB_PEP.*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB_PEP.*
7: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB_PEP.*
8: /cgn2_6/ptodata/1/pubppaa/US09_PUB_PEP.*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB_PEP.*
10: /cgn2_6/ptodata/1/pubppaa/US05_PUBCOMB_PEP.*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB_PEP.*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB_PEP.*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB_PEP.*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	130	100.0	24	US-09-785-059-10	Sequence 10, Appl
2	130	100.0	24	US-10-079-075-10	Sequence 10, Appl
3	130	100.0	24	US-09-785-058-10	Sequence 10, Appl
4	130	100.0	36	US-09-785-059-11	Sequence 11, Appl
5	130	100.0	36	US-10-079-075-11	Sequence 11, Appl
6	130	100.0	36	US-09-785-058-11	Sequence 11, Appl
7	130	100.0	48	US-09-785-059-12	Sequence 12, Appl
8	130	100.0	48	US-10-079-075-12	Sequence 12, Appl
9	130	100.0	24	US-09-785-058-12	Sequence 12, Appl
10	88	67.7	24	US-10-079-075-5	Sequence 5, Appl
11	88	67.7	24	US-10-079-075-5	Sequence 5, Appl
12	88	67.7	24	US-09-785-058-5	Sequence 5, Appl
13	88	67.7	36	US-09-785-059-6	Sequence 6, Appl
14	88	67.7	36	US-10-079-075-6	Sequence 6, Appl
15	88	67.7	36	US-09-785-058-6	Sequence 6, Appl
16	88	67.7	42	US-09-785-059-7	Sequence 7, Appl
17	88	67.7	42	US-10-079-075-7	Sequence 7, Appl
18	88	67.7	42	US-09-785-058-7	Sequence 7, Appl
19	88	67.7	48	US-09-785-059-8	Sequence 8, Appl

20	88	67.7	48	9	US-10-079-075-8	Sequence 8, Appl
21	88	67.7	48	9	US-09-785-058-8	Sequence 8, Appl
22	78	60.0	31	9	US-09-785-059-3	Sequence 3, Appl
23	78	60.0	31	9	US-10-079-075-3	Sequence 3, Appl
24	78	60.0	31	9	US-09-785-058-3	Sequence 3, Appl
25	61	46.9	12	9	US-09-785-059-9	Sequence 9, Appl
26	61	46.9	12	9	US-10-079-075-9	Sequence 9, Appl
27	61	46.9	12	9	US-09-785-058-9	Sequence 9, Appl
28	57	43.8	394	10	US-10-081-816-40	Sequence 10, Appl
29	56	43.1	634	10	US-09-733-665-10	Sequence 10, Appl
30	56	43.1	635	10	US-09-733-665-8	Sequence 8, Appl
31	56	43.1	635	10	US-09-733-665-9	Sequence 9, Appl
32	52.5	40.4	768	10	US-09-733-665-6	Sequence 6, Appl
33	52.5	40.4	831	10	US-09-733-665-7	Sequence 7, Appl
34	52	40.0	770	10	US-09-992-896-9	Sequence 9, Appl
35	52	40.0	770	10	US-09-815-656-31	Sequence 31, Appl
36	51	39.2	243	9	US-09-738-626-6003	Sequence 6003, Ap
37	50.5	38.8	19	10	US-09-864-761-38480	Sequence 38480, A
38	47.5	36.5	219	9	US-09-738-626-3670	Sequence 3670, Ap
39	47	36.2	12	9	US-09-785-059-4	Sequence 4, Appl
40	47	36.2	12	9	US-10-079-075-4	Sequence 4, Appl
41	47	36.2	12	9	US-09-785-058-4	Sequence 4, Appl
42	47	36.2	30	10	US-09-864-761-47840	Sequence 47840, A
43	47	36.2	31	9	US-09-785-059-2	Sequence 2, Appl
44	47	36.2	31	9	US-10-079-075-2	Sequence 2, Appl
45	47	36.2	31	9	US-09-785-058-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-785-059-10
; Sequence 10, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mielzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-10

Query Match 100.0%; Score 130; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRMVRRVRRVRRVRRVRRVRR 24
DB 1 RRMVRRVRRVRRVRRVRRVRR 24

RESULT 2
US-10-079-075-10
; Sequence 10, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mielzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12

XX The present invention relates to novel peptides containing a sequence of
 CC 18 amino acids with either a hydrophobic surface of 5-7 amino acids and
 CC 80% hydrophobic amino acids or a hydrophobic surface of 5-6 amino acids,
 CC 80% hydrophobic amino acids, and 50% Arg and/or Lys, and a second
 CC hydrophobic surface of 2-4 hydrophobic amino acids or a second
 CC hydrophobic surface of 3-5 hydrophobic amino acids and 80% hydrophobic
 CC amino acids. The peptides of the invention are used for the effective
 CC introduction of a desired gene in gene therapy, especially for treatment
 CC and prevention of graft-versus-host disease (GVHD) and restenosis
 CC following cancer or arteriosclerosis. The present amino acid sequence
 CC represents one of a collection (AAU9952-AAU9978) of synthetic peptides
 CC of the invention which have an affinity with phosphatidylserine.

XX SQ Sequence 37 AA;

Query Match 42.3%; Score 55; DB 23; Length 37;
 Best Local Similarity 50.0%; Pred. No. 2.6;
 Matches 12; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy 1 RRMVRRVRVRVRVRVRVRVR 24
 |||||:|||||
 Db 9 RRMVRRRLRLRLRLRHRVRVR 32

RESULT 14

AAW99084
 ID AAW99084 standard; Protein; 761 AA.

XX AAW99084;

DT 13-MAY-1999 (first entry)

DE Non-B, non-C, non-G hepatitis virus protein sequence SEQ ID NO:63.

XX Hepatitis virus; non-B non-C non-G hepatitis virus; PCR primer;

KM virion; detection; screening; infection; blood transfusion.

XX Hepatitis virus.

XX MO9905282-A1.

PD 04-FEB-1999.

XX 27-JUL-1998; 98WO-JP03340.

XX 13-MAR-1998; 98JP-0082962.

PR 25-JUL-1997; 97JP-0233246.

PR 09-OCT-1997; 97JP-0314196.

XX (TAMU/) TAMURA R.

PI Nishizawa T, Okamoto H;

DR WPI, 1999-142937/12.

PT Gene isolated from non-B non-C non-G hepatitis DNA virus - and its
 PT expression products, useful for diagnosis and treatment of hepatitis
 PT and screening of blood for transfusion.

XX Claim 46; Page 104-106; 113pp; Japanese.

XX The present invention describes a gene which is contained in a non-B
 CC non-C non-G hepatitis DNA virus. It is 3500-4000 bases in length and
 CC contains two overlapping open reading frames (ORF). It is obtained by
 CC polymerase chain reaction (PCR) amplification using primer. The present
 CC sequence represents a specifically claimed non-B non-C non-G hepatitis
 CC protein sequence. The gene can be used for the production of vaccines
 CC for prevention and treatment of non-B non-C non-G hepatitis infection.
 CC Diagnosis of such infection, and screening of blood (e.g. intended for
 CC transfusion) for the presence of the virus, by using the virion or
 CC antigenic peptides as reagents for detection of antibodies to the virus,
 CC or by direct detection of the gene using PCR with primers derived from

CC the gene sequence.

XX SQ Sequence 761 AA;

Query Match 42.3%; Score 55; DB 20; Length 761;
 Best Local Similarity 58.3%; Pred. No. 45;
 Matches 14; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 RRMVRRVRVRVRVRVRVRVR 24
 |||||:|||||
 Db 6 RRMVRRVRVRVRVRVRVRVR 29

RESULT 15

AAW99084
 ID AAW99084 standard; protein; 120 AA.

XX AAW99084;

DT 20-OCT-2000 (first entry)

DE Protein fragment #1 used in TT virus antibody determination.

XX Serum type classification; TT virus; antibody; viral infection;

KM identify; treatment.

XX TT virus.

XX JP2000135087-A.

PD 16-MAY-2000.

XX 29-OCT-1998; 98JP-0309208.

PR 29-OCT-1998; 98JP-0309208.

XX (SRLS-) SRL KK.

XX WPI, 2000-415430/36.

PT Peptides for determination of anti-TT virus antibody and method for
 PT serum classification of TT virus using the peptides

XX Claim 1; Page 6-7; 12pp; Japanese.

XX A method for serum type classification of TT virus (also known as
 CC hepatitis TT virus) has been identified. The method relies on the use of
 CC peptide fragments of the virus. The invention also relates to the use of
 CC TT virus peptides for anti-TT virus antibody determination. The anti-TT
 CC virus antibodies and the serum type classification method, can be used to
 CC screen TT virus, to determine its route of infection, and seroconversion.
 CC The classification of TT virus may lead to improved treatment of viral
 CC disease. The present sequence represents a fragment of TT virus protein
 CC used in the course of the invention.

XX SQ Sequence 120 AA;

Query Match 41.5%; Score 54; DB 21; Length 120;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 RRMVRRVRVRVRVRVRVR 20
 |||||:|||||
 Db 11 RRMVRRVRVRVRVRVRVR 30

Search completed: June 9, 2003, 11:55:36
 Job time : 26.8723 secs

```
XX
AC AAB84456;
XX
DT 22-AUG-2001 (first entry)
DE Amino acid sequence of a Sentinel virus I (SVI) protein.
XX
KM SVI; viral replication; viral infection; vaccine.
XX
OS Sentinel virus I.
XX
PN WO200142299-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-IB02011.
XX
PR 10-DEC-1999; 99US-0172696.
XX
PA (HOFF ) ROCHE DIAGNOSTICS GMBH.
XX
PI Liu J, Bohenzky RA, Lin Y, Chen BP;
XX
DR WPI; 2001-381643/40.
XX
PT Novel virus, designated sentinel virus I, associated with cryptogenic,
PT nona-G hepatitis, and polynucleotides and polypeptides of virus useful
PT for detecting SVI virus and/or SVI virus infection.
XX
PS Example 1; Page 58-59; 65pp; English.
XX
CC The present sequence represents a protein of Sentinel virus I (SVI).
CC SVI polynucleotides are useful for detecting SVI virus. Probes
CC and primers derived from SVI polynucleotide sequences are useful for
CC identifying and isolating new variants of SVI. SVI polynucleotides are
CC useful for detecting SVI virus, producing SVI polypeptides, constructing
CC SVI-based expression/transduction vectors and as antisense
CC oligonucleotides or for construction of antisense SVI vectors. Antisense
CC SVI polynucleotides block expression of SVI proteins and/or SVI viral
CC replication in SVI infected cells, and thus are useful for treating SVI
CC infections. SVI polypeptides are useful in vaccines for preventing SVI
CC infection and for treating SVI infection.
XX
SQ Sequence 635 AA;
XX
Query Match 43.1%; Score 56; DB 22; Length 635;
Best Local Similarity 45.7%; Pred. No. 29;
Matches 16; Conservative 1; Mismatches 4; Indels 14; Gaps 3;
OY 1 RRVVRRVRRVRRVRRV-----RRVRR 24
Db 20 RRV-RRLR--WRRPRRAVRRRRRGRVRRRRMAR 51
XX
RESULT 12
AAB84457
ID AAB84457 standard; Protein; 635 AA.
XX
AC AAB84457;
XX
DT 22-AUG-2001 (first entry)
DE Amino acid sequence of a Sentinel virus I (SVI) protein.
XX
KM SVI; viral replication; viral infection; vaccine.
XX
OS Sentinel virus I.
XX
PN WO200142299-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-IB02011.
```

```
XX
PR 10-DEC-1999; 99US-0172696.
XX
PA (HOFF ) ROCHE DIAGNOSTICS GMBH.
XX
PI Liu J, Bohenzky RA, Lin Y, Chen BP;
XX
DR WPI; 2001-381643/40.
XX
PT Novel virus, designated sentinel virus I, associated with cryptogenic,
PT nona-G hepatitis, and polynucleotides and polypeptides of virus useful
PT for detecting SVI virus and/or SVI virus infection.
XX
PS Example 1; Page 60-62; 65pp; English.
XX
CC The present sequence represents a protein of Sentinel virus I (SVI).
CC SVI polynucleotides are useful for detecting SVI virus. Probes
CC and primers derived from SVI polynucleotide sequences are useful for
CC identifying and isolating new variants of SVI. SVI polynucleotides are
CC useful for detecting SVI virus, producing SVI polypeptides, constructing
CC SVI-based expression/transduction vectors and as antisense
CC oligonucleotides or for construction of antisense SVI vectors. Antisense
CC SVI polynucleotides block expression of SVI proteins and/or SVI viral
CC replication in SVI infected cells, and thus are useful for treating SVI
CC infections. SVI polypeptides are useful in vaccines for preventing SVI
CC infection and for treating SVI infection.
XX
SQ Sequence 635 AA;
XX
Query Match 43.1%; Score 56; DB 22; Length 635;
Best Local Similarity 45.7%; Pred. No. 29;
Matches 16; Conservative 1; Mismatches 4; Indels 14; Gaps 3;
OY 1 RRVVRRVRRVRRVRRV-----RRVRR 24
Db 20 RRV-RRLR--WRRPRRAVRRRRRGRVRRRRMAR 51
XX
RESULT 13
AAU79973
ID AAU79973 standard; Peptide; 37 AA.
XX
AC AAU79973;
XX
DT 15-JUL-2002 (first entry)
DE Phosphatidylserine affinity synthetic peptide #22.
XX
KM Phosphatidylserine; graft-versus-host-disease; GVHD; restenosis; cancer;
KM arteriosclerosis; immunosuppressive; vasotropic; gene therapy; tumour;
XX
XX peptide therapy.
XX
OS Synthetic.
XX
PN WO200230961-A1.
XX
PD 18-APR-2002.
XX
PF 28-SEP-2001; 2001WO-JP08565.
XX
PR 12-OCT-2000; 2000JP-0312600.
XX
PA (MOCH ) MOCHIDA PHARM CO LTD.
XX
PI Kuriyama S, Taguchi Y;
XX
DR WPI; 2002-362674/39.
XX
PT New peptides that bind to a gene and improve the transfer of the gene
PT into cells for use in a high efficiency gene therapy vector, for
PT treating and preventing of graft-versus-host-disease and restenosis -
XX
XX Claim 7; Page 136; 146pp; Japanese.
```


PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147203.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.2%; Score 57.5; DB 21; Length 95;
Best Local Similarity 56.0%; Pred. No. 3.1;
Matches 14; Conservative. 1; Mismatches 9; Indels 1; Gaps 1;

Qy 1 RRVRRRRRVR-RVVRVVRRVR 24
Db 50 RRVRTLRRRWRKRRRTLRRRWR 74

RESULT 6
ABB90000
ID ABB90000 standard; Protein; 136 AA.
XX AC ABB90000;

DT 24-MAY-2002 (first entry)
XX XX

DE Human polypeptide SEQ ID NO 2376.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antidiabetic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.

XX Homo sapiens.

OS Homo sapiens.
PN WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US16450.

XX 19-MAY-2000; 2000US-205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

PI

PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158233.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

```

Query Match      44.2%   Score 57.5; DB 21; Length 82;
Best Local Similarity 56.5%; Pred. No.2.7;
Matches 13; Conservative 2; Mismatches 3; Indels 5; Gaps 2

QY          1 RRVRRVRRVRRVVRRVVRRVR 23
             ||| | :|| ||| 
Db           50 RRW-RTLRRWRMR-----IRRWRR 67

RESULT 5
AAGS0720
ID AAGS0720 standard; Protein; 95 AA.
XX
XX AAGS0720;
XX AC
XX DT 18-OCT-2000 (first entry)
XX DE
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX KM
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX PN
XX PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX
```

[illegible]

XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

PT Interactions --
XX Disclosure; SEQ ID NO 10746; 21pp + Sequence Listing; English.
XX
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB161840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 133 AA;

Query Match 50.8%; Score 66; DB 22; Length 133;
Best Local Similarity 65.0%; Pred. No. 0.39;
Matches 13; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 2 RWRVRVRVRVRVRVRVRVR 21
DB 105 RWRVRVRVRVRVRVRVR 122

RESULT 2
AAB03840
ID AAB03840 standard; protein; 120 AA.
XX
XX AAB03840;
AC
XX
XX
DT 20-OCT-2000 (first entry)
XX
XX Protein fragment #2 used in TT virus antibody determination.
DE
XX Serum type classification; TT virus; antibody; viral infection;
KM identify; treatment.
XX
XX TT virus.
OS
XX JP2000135087-A.
PN
XX
XX 16-MAY-2000.
PD
XX
XX 29-OCT-1998; 98UP-0309208.
PF
XX
XX 29-OCT-1998; 98UP-0309208.
PR
XX 29-OCT-1998; 98UP-0309208.
XX
XX (SRLS-) SRL KK.
PA
XX
XX WPI; 2000-415430/36.
DR
XX
XX
XX Peptides for determination of anti-TT virus antibody and method for
PT serum classification of TT virus using the peptides
XX
XX
PS Claim 1; Page 7; 12pp; Japanese.
XX
XX A method for serum type classification of TT virus (also known as
CC hepatitis TT virus) has been identified. The method relies on the use of
CC peptide fragments of the virus. The invention also relates to the use of
CC TT virus peptides for anti-TT virus antibody determination. The anti-TT
CC virus antibodies and the serum type classification method, can be used to
CC screen TT virus, to determine its route of infection, and seroconversion.
CC The classification of TT virus may lead to improved treatment of viral
CC disease. The present sequence represents a fragment of TT virus protein
CC used in the course of the invention.
XX
SQ Sequence 120 AA;

Query Match 46.2%; Score 60; DB 21; Length 120;
Best Local Similarity 63.6%; Pred. No. 1.9;

Matches 14; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 3 WWRVRVRVRVRVRVRVRVR 24
DB 6 WWRVRVRVRVRVRVRVR 25

RESULT 3
AAM47987
ID AAM47987 standard; protein; 765 AA.
XX
XX AAM47987;
AC
XX
XX 07-MAR-2002 (first entry)
DT
XX
XX Simian TTV CH65-1 SEQ ID NO 2.
DE
XX
XX Simian TTV; virus; diagnosis; hepatitis; CH65-1.
KM
XX
XX Simian TTV; virus; diagnosis; hepatitis; CH65-1.
XX
XX Simian TT virus.
OS
XX
XX WO200185771-A1.
PN
XX
XX 15-NOV-2001.
PD
XX
XX 11-MAY-2001; 2001WO-JP03954.
PF
XX
XX 11-MAY-2000; 2000JP-0137894.
PR
XX 08-NOV-2000; 2000JP-0340614.
XX
XX (EISA) EISAI CO LTD.
XX
XX
XX Abe K;
PI
XX
XX WPI; 2002-097552/13.
DR
XX
XX N-PSDB; ABA05987.
DE
XX
XX DNA and proteins for diagnosis of non-A, non-B, non-C hepatitis
PT
XX
XX Disclosure; Page 20-22; 37pp; Japanese.
XX
XX
XX The invention relates to an isolated nucleic acid, comprising a 3899
CC (S1-ABA05987) or 3322 (S2-ABA05995) nucleotide sequence fully defined in
CC the specification useful for diagnosis of non-A, non-B, non-C hepatitis.
CC The present sequence is that of a TTV polypeptide, useful to the
CC invention.
XX
XX
SQ Sequence 765 AA;

Query Match 45.8%; Score 59.5; DB 23; Length 765;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 15; Conservative 0; Mismatches 4; Indels 5; Gaps 2;

QY 1 RWRVRVRVRVRVRVRVRVR 24
DB 7 RWRVRVRVRVRVRVRVR 25

RESULT 4
AAG28909
ID AAG28909 standard; protein; 82 AA.
XX
XX AAG28909;
AC
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 34302.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:35:06 ; Search time 25.8723 Seconds
(without alignments)
123.607 Million cell updates/sec

Title: US-10-079-075-10

Sequence: 1 RRVRRVRRVRRVRRVRR 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

1. A_geneseq_101002.*
2. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1980.DAT.*
3. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1981.DAT.*
4. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1982.DAT.*
5. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1983.DAT.*
6. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1984.DAT.*
7. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1985.DAT.*
8. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1986.DAT.*
9. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1987.DAT.*
10. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1988.DAT.*
11. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1990.DAT.*
12. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1991.DAT.*
13. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1992.DAT.*
14. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1993.DAT.*
15. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1994.DAT.*
16. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1995.DAT.*
17. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1996.DAT.*
18. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1997.DAT.*
19. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1998.DAT.*
20. /SIDS2/gcgdata/geneseq/genseq-emb1/AA1999.DAT.*
21. /SIDS2/gcgdata/geneseq/genseq-emb1/AA2000.DAT.*
22. /SIDS2/gcgdata/geneseq/genseq-emb1/AA2001.DAT.*
23. /SIDS2/gcgdata/geneseq/genseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	66	50.8	133	22	ABB61318	Drosophila melanog
2	60	46.2	120	21	AAB03840	Protein fragment #
3	59.5	45.8	765	23	AAW47987	Simian TTV CH65-1
4	57.5	44.2	82	21	AAQ28909	Arabidopsis thalia
5	57.5	44.2	95	21	AAQ50720	Arabidopsis thalia
6	57	43.8	136	23	AAB90000	Human polypeptide
7	57	43.8	337	22	AAB75822	Human polypeptide
8	57	43.8	643	22	ABG22851	Drosophila gustatc
9	57	43.8	681	22	ABB68280	Novel human diagno
10	56	43.1	634	22	ABB84458	Drosophila melanog
						Amino acid sequenc

11	56	43.1	635	22	AA884456	Amino acid sequenc
12	56	43.1	635	22	AA884457	Amino acid sequenc
13	55	42.3	37	23	AAU79973	Phosphatidylserine
14	55	42.3	761	20	AAW99084	Non-B, non-C, non-
15	54	41.5	120	21	AA803839	Protein fragment #
16	53.5	41.2	743	21	AA811535	SEN virus protein
17	52.5	40.4	147	21	AA655794	Arabidopsis thalia
18	52.5	40.4	727	23	AA847969	Simian TTV CH71 SB
19	52.5	40.4	768	22	AA844544	Amino acid sequenc
20	52.5	40.4	831	22	AA884455	Amino acid sequenc
21	52	40.0	18	22	AA655539	Pepide sequence u
22	52	40.0	30	21	AA803851	Pepide P0130 used
23	52	40.0	756	20	AA847795	TT virus clone X94
24	52	40.0	770	23	AAW99082	Non-B, non-C, non-
25	52	40.0	770	21	AAV97119	TT virus (TTV-GH1)
26	52	40.0	770	21	AA837531	TT virus ORF1 prot
27	51.5	39.6	105	21	AA654778	Arabidopsis thalia
28	51	39.2	243	22	AA692249	C glutamicum prote
29	51	39.2	243	22	AA876634	Corynebacterium gl
30	51	39.2	388	20	AAV04998	Mycobacterium spec
31	51	39.2	762	21	AA811546	SEN virus protein
32	51	39.2	898	18	AAW31853	Mycobacterium tube
33	50.5	38.8	19	22	AA837947	Pepide #5433 enco
34	50.5	38.8	19	22	AA823182	Pepide #5181 enco
35	50.5	38.8	19	22	AAW85859	Human brain expres
36	50.5	38.8	19	22	AAW71060	Human bone marrow
37	50.5	38.8	19	22	AAW18821	Pepide #5285 enco
38	50.5	38.8	19	22	AAW31337	Pepide #5374 enco
39	50.5	38.8	19	23	AA6040853	Human peptide enco
40	50	38.5	28	10	AA691336	Novel amino acid sequenc
41	50	38.5	93	22	AA632962	Novel human diagno
42	50	38.5	106	20	AAV44835	Chlamydia pneumoni
43	50	38.5	130	22	AA630960	Novel human diagno
44	50	38.5	201	22	AA637999	Novel human diagno
45	49.5	38.1	76	22	AAU52184	Propionibacterium

ALIGNMENTS

XX	ABRB61318	standard; Protein; 133 AA.
XX	ABRB61318;	
XX	26-MAR-2002	(first entry)
DE	Drosophila melanogaster	polypeptide SEQ ID NO 10746.
XX	Drosophila; developmental biology; cell signalling; insecticide;	
XX	pharmaceutical.	
OS	Drosophila melanogaster.	
XX	WO200171042-A2.	
XX	27-SEP-2001.	
XX	23-MAR-2001; 2001WO-US09231.	
XX	23-MAR-2000; 2000US-191637P.	
PR	11-JUL-2000; 2000US-0614150.	
XX	(PEKE) PE CORP NY.	
XX	Venter JC, Adams M, Li PWD, Myers EW,	
XX	WPI; 2001-556660/75.	
DR	N-PSDB; ABL05421.	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more	
PT	genes from Drosophila and for elucidating cell signalling and cell-cell	

DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE KIAA0278 protein (HARG) (Activity-regulated cytoskeleton-associated
protein) (Fragment).
GN KIAA0278 OR HARG OR ARC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRNIN;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawabayasi Y., Ohara O.,
Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
the coding sequences of 80 new genes (K1A0201-K1A0280) deduced by
analysis of cDNA clones from cell line KG-1 and brain."
RL DNA Res. 3:321-329(1996).
RN [2]
RP SEQUENCE OF 65-460 FROM N.A.
RC TISSUE=FOREBRNIN; Wendholt D., Liedtke T., Barnekow A.;
RA Kremerskothen J., Barnekow A.;
RT "Harg, a gene highly homologous to the rat arg3.1 gene."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 65-460 FROM N.A.
RX MEDLINE=21003425; PubMed=11117363;
RA Kremerskothen J., Barnekow A.;
RT "Human activity-regulated cytoskeleton-associated gene (ARC) maps to
chromosome 8q24."
RL Chromosome Res. 8:655-655(2000).
RN [4]
RP SEQUENCE OF 65-460 FROM N.A.
RX MEDLINE=20429191; PubMed=10970730;
RA Haug K., Kremerskothen J., Hallmann K., Sander T., Dullinger J.,
Rau B., Beyenburg S., Lentze M.J., Barnekow A., Elger C.E.,
RA Proping P., Heils A.;
RT "Mutation screening of the chromosome 8q24.3-human activity-regulated
cytoskeleton-associated gene (ARC) in idiopathic generalized
epilepsy."
RL Mol. Cell. Probes 14:255-260(2000).
RN [5]
RP SEQUENCE OF 65-460 FROM N.A.
RC TISSUE=UTERUS;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; D87468; BAI19667.1; -.
DR EMBL; AF193421; AAF07185.1; -.
DR EMBL; AF248637; AAC33705.1; -.
DR EMBL; BC012321; AAH12321.1; -.
PT NON_TER 1
SQ SEQUENCE 460 AA; 52113 MW; AD6A0A3905E156CE CRC64;
QY
Query Match 47.7%; Score 62; DB 4; Length 460;
Best Local Similarity 47.8%; Pred. No. 2.1;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
Db 171 RRVRRVRRVRRVRRVRRVRR 193
QY 2 RRVRRVRRVRRVRRVRRVRR 24
ID Q9DUB7 PRELIMINARY; PRT; 720 AA.
AC Q9DUB7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE ORF1.
OS TT virus.
OS Viruses; ssDNA viruses; unclassified ssDNA viruses.

OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AT-TTV3;
RL Okamoto H.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AT-TTV3;
RX MEDLINE=20534983; PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
phylogenetic relatedness."
RL Virology 277:368-378(2000).
DR EMBL; AB041961; BAB19320.1; -.
DR InterPro; IPR004219; TTVvirus_Unk.
DR Pfam; PF02956; TT_ORF1; 1.
SQ SEQUENCE 720 AA; 85948 MW; 4915FE9AEB503E15 CRC64;
QY
Query Match 47.7%; Score 62; DB 12; Length 720;
Best Local Similarity 64.0%; Pred. No. 3.2;
Matches 16; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
Db 5 RRVRRVRRVRRVRRVRRVRR 26
QY 1 RRVRRVRRVRRVRRVRRVRR 24
ID Q9DUB8 PRELIMINARY; PRT; 767 AA.
AC Q9DUB8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE Hypothetical 90.4 kDa protein.
OS TT virus.
OS Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=US35;
RX MEDLINE=99350006; PubMed=10423143;
RA Ecker J.C., Leary T.P., Desai S.M., Chalmers M.L., Mushahwar I.K.;
RT "Analyses of TT virus full-length genomic sequences."
RL J. Gen. Virol. 80:1743-1750(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=US35;
RA Ecker J.C., Leary T.P., Desai S.M., Chalmers M.L., Mushahwar I.K.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF122820; AAD45650.1; -.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR004219; TTVvirus_Unk.
DR Pfam; PF02956; TT_ORF1; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 767 AA; 90383 MW; 6FAC51C088C1E10 CRC64;
QY
Query Match 47.7%; Score 62; DB 12; Length 767;
Best Local Similarity 62.5%; Pred. No. 3.4;
Matches 15; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
Db 4 RRVRRVRRVRRVRRVRRVRR 25
QY 1 RRVRRVRRVRRVRRVRRVRR 24
ID Q8V7C3 PRELIMINARY; PRT; 49 AA.
AC Q8V7C3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE ORF1.
OS TT virus.
OS Viruses; ssDNA viruses; unclassified ssDNA viruses.

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=CBA X C57BL/6;
RA Medrano S., Worley P.F., Chowdhury S., Lananan A., Steward O.,
RA Scrable H.;
RT "Characterization of the promoter region of the immediate early gene
ARC.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF254662; AAG10254.1; -.
FT NON_TER 242 242
SQ SEQUENCE 242 AA; 27428 MW; C28034A022C30D20 CRC64;

Query Match 47.7%; Score 62; DB 11; Length 242;
Best Local Similarity 47.8%; Pred. No. 1.1;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 2 RWVRVRVRVRVRVRVRVRVR 24
DB 107 RWVKREHVRVREVFYRLERWADR 129

RESULT 6
ID Q62743 PRELIMINARY; PRT; 396 AA.
AC Q62743;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE Growth factor.
GN ARC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RC MEDLINE=95161073; PubMed=7857651;
RA Lyford G.L., Yamagata K., Kaufmann W.E., Barnes C.A., Sanders L.K.,
RA Copeland N.G., Gilbert D.O., Jenkins N.A., Lananan A.A., Worley P.F.;
RT "Arc, a growth factor and activity-regulated gene, encodes a novel
cytoskeleton-associated protein that is enriched in neuronal
dendrites.";
RL Neuron 14:433-445(1995).
DR EMBL: U19866; AAA68695.1; -.
SQ SEQUENCE 396 AA; 45367 MW; D975A1848F8AC94A CRC64;

Query Match 47.7%; Score 62; DB 11; Length 396;
Best Local Similarity 47.8%; Pred. No. 1.8;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 2 RWVRVRVRVRVRVRVRVRVR 24
DB 107 RWVKREHVRVREVFYRLERWADR 129

RESULT 7
ID Q63053 PRELIMINARY; PRT; 396 AA.
AC Q63053;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE Arc3.1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.

```

```

RC STRAIN=SPRAGUE DAWLEY, TISSUE=HIPPOCAMPUS;
RX MEDLINE=95296386; PubMed=7777577;
RA Link W., Konieczko U., Kauselmann G., Krug M., Schwanke B., Frey U.,
RA Kuhl D.;
RT "Somatodendritic expression of an immediate early gene is regulated by
RT synaptic activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5734-5738(1995).
DR EMBL: Z46925; CAA87033.1; -.
SQ SEQUENCE 396 AA; 45353 MW; 4E95B46B75853CA5 CRC64;

Query Match 47.7%; Score 62; DB 11; Length 396;
Best Local Similarity 47.8%; Pred. No. 1.8;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 2 RWVRVRVRVRVRVRVRVRVR 24
DB 107 RWVKREHVRVREVFYRLERWADR 129

RESULT 8
ID Q9WV31 PRELIMINARY; PRT; 396 AA.
AC Q9WV31;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Growth factor ARC (ARC3.1/ARC) (Activity regulated
DE cytoskeletal-associated protein).
GN ARC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=FOREBRAIN;
RA Chowdhury S., Lananan A.A., Worley P.F.;
RT "The Arc gene, a mouse homolog of rat Arc.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=129/SV(EV);
RX MEDLINE=21359664; PubMed=11466419;
RA Waltereit R., Dammernann B., Wulff P., Scaifidi J., Staubli U.,
RA Kauselmann G., Bundman M., Kuhl D.;
RT "Arc3.1/Arc mRNA Induction by Ca2+ and cAMP Requires Protein Kinase A
RT and Mitogen-Activated Protein Kinase/Extracellular Regulated Kinase
RT Activation.";
RL J. Neurosci. 21:5484-5493(2001).
RN [3]
RC SEQUENCE FROM N.A.
RC Strausberg R.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF162777; AAD43586.1; -.
DR EMBL: AF177701; AAK91587.1; -.
DR EMBL: BC023127; AAK23127.1; -.
DR MGD: MGI:88067; Arc.
SQ SEQUENCE 396 AA; 45321 MW; F4D3505FDA77D18A CRC64;

Query Match 47.7%; Score 62; DB 11; Length 396;
Best Local Similarity 47.8%; Pred. No. 1.8;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 2 RWVRVRVRVRVRVRVRVRVR 24
DB 107 RWVKREHVRVREVFYRLERWADR 129

RESULT 9
ID Q9Y469 PRELIMINARY; PRT; 460 AA.
AC Q9Y469;
DT 01-NOV-1999 (TREMBlrel. 12, Created)

```


RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matvei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Stirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RT "The genome sequence of Dirosophila melanogaster.",
 RL Science 287:2185-2195(2000).
 DR EMBL, AB003504; AAF48711.1; -.
 DR FlyBase; FBgn0030830; CG5172.
 SQ SEQUENCE 133 AA; 14510 MW; 161861BFE8ACB6C6 CRC64;

Query Match 50.8%; Score 66; DB 5; Length 133;
 Best Local Similarity 65.0%; Pred. No. 0.19;
 Matches 13; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

2 RRVRRVRRVRRVRRVRRV 21
 105 RWRRTTRRLMRR-RSMRR 122

RESULT 2

ID Q91FV0 PRELIMINARY; PRT; 766 AA.

AC Q91FV0;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative capsid protein.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 ON NCBI_TaxID=68887;

RA MEDLINE=20409089; PubMed=10950985;
 RA Hallett R.L., Clewley J.P., Bobet F., McKiernan P.J., Teo C.G.;
 RT "Characterization of a highly divergent TT virus genome."; [2]
 RL J. Gen. Virol. 81:2273-2279(2000).

RP SEQUENCE FROM N.A.
 RC STRAIN=PMV;
 RA Hallett R.L., Clewley J.P., Teo C.G.;
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL, AAF61761; AAF82559.1; -.
 DR InterPro; IPR004219; TTVirus_unk.
 DR Pfam; PF02956; TT_ORF1.1.

SQ SEQUENCE 766 AA; 90844 MW; 2255A3BBA3CC8187 CRC64;
 Query Match 49.2%; Score 64; DB 12; Length 766;
 Best Local Similarity 42.1%; Pred. No. 1.9;
 Matches 16; Conservative 0; Mismatches 8; Indels 14; Gaps 1;

1 RRVRRVRRVRRVRRVRRV 24
 11 RWRRTTRRLMRR-RSMRR 48

RESULT 3

ID Q9OU36 PRELIMINARY; PRT; 656 AA.

AC Q9OU36;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

6

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA, complete genome, isolate:TLMV-CBD203.
 OS TTV-like mini virus.
 OC Viruses; ssDNA viruses; Circoviridae.
 ON NCBI_TaxID=93678;

RP SEQUENCE FROM N.A.
 RC STRAIN=TLMV-CBD203;
 RA Mishiro S.;
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.

RP SEQUENCE FROM N.A.
 RC STRAIN=TLMV-CBD203;
 RA MEDLINE=20338593; PubMed=10881684;
 RA Takahashi K., Iwasa Y., Hijioka M., Mishiro S.;
 RT "Identification of a new human DNA virus (TTV-like mini virus, TTMV) intermediate related to TT virus and chicken anemia virus."; Arch. Virol. 145:979-993(2000).
 DR EMBL, AB026929; BAA86945.1; -.
 DR InterPro; IPR004219; TTVirus_unk.
 DR Pfam; PF02956; TT_ORF1.1.

SQ SEQUENCE 656 AA; 76983 MW; DA2CA1D3C2D83A37 CRC64;

Query Match 48.5%; Score 63; DB 12; Length 656;
 Best Local Similarity 47.1%; Pred. No. 2.2;
 Matches 16; Conservative 1; Mismatches 7; Indels 10; Gaps 2;

1 RRVRRVRRVRRVRRVRRV 24
 6 RWRRTTRRLMRR-RSMRR 39

RESULT 4

ID Q91CZ2 PRELIMINARY; PRT; 759 AA.

AC Q91CZ2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ORF1.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 ON NCBI_TaxID=68887;

RP SEQUENCE FROM N.A.
 RX MEDLINE=21488921; PubMed=11601907;
 RA Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
 RA Yoshikawa A.;
 RT "Heterogeneous distribution of TT virus of distinct genotypes in multiple tissues from infected humans."; J. Virol. 75:358-368(2001).
 RL Virology 288:358-368(2001).
 DR EMBL, AB060595; BAB69912.1; -.
 DR InterPro; IPR004219; TTVirus_unk.
 DR Pfam; PF02956; TT_ORF1.1.

SQ SEQUENCE 759 AA; 89860 MW; 5B79DFD71A37010D CRC64;

Query Match 48.1%; Score 62.5; DB 12; Length 759;
 Best Local Similarity 47.1%; Pred. No. 2.9;
 Matches 17; Conservative 0; Mismatches 7; Indels 11; Gaps 1;

1 RRVRRVRRVRRVRRVRRV 24
 19 RWRRTTRRLMRR-RSMRR 39

RESULT 5

ID Q9ES15 PRELIMINARY; PRT; 242 AA.

AC Q9ES15;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE ARC (Fragment).


```
CC SEQUENCE FROM N.A.
CC SPECIES=Phage HK022;
CC MEDLINE=93342457; PubMed=2547971;
CC Yagil E., Dolev S., Oberto J., Kislay N., Ramalah N., Weisberg R.A.;
CC "determination of site-specific recombination in the lambdaoid
CC coliphage HK022. An evolutionary change in specificity.";
CC J. Mol. Biol. 207:695-717(1989).
CC -I- FUNCTION: EXCISIONASE AND INTEGRASE ARE NECESSARY FOR THE EXCISION
CC OF PROPHAGE FROM THE HOST GENOME BY SITE-SPECIFIC RECOMBINATION AT
CC THE ATT SITE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M60848; AAA67901.1; -
CC DR EMBL; X51962; CAA36222.1; -
CC DR PIR; S06533; S06533.
CC KM DNA recombination; DNA-binding.
CC SQ SEQUENCE 72 AA; 8635 MW; 0E6A4843503344AA CRC64;

Query Match 35.4%; Score 46; DB 1; Length 72;
Best Local Similarity 47.8%; Pred. No. 2.6;
Matches 11; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 1 RRVVRRVRRVRRVRRVRRVRR 23
Db 6 QEWNAQR--PRLSTVRRVRR 26

RESULT 14
VXIS_LAMB
ID VXIS_LAMB STANDARD; PRT; 72 AA.
AC P03699;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Excisionase.
GN XIS.
OS Bacteriophage lambda.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83189071; PubMed=622115;
RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;
RT "Nucleotide sequence of bacteriophage lambda DNA.";
RL J. Mol. Biol. 162:729-773(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81053845; PubMed=6253947;
RA Davies R.W.;
RT "DNA sequence of the int-xis-pi region of the bacteriophage lambda;
RT overlap of the int and xis genes.";
RL Nucleic Acids Res. 8:1765-1782(1980).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=80234646; PubMed=6446713;
RA Hoese R.H., Foeller C., Bidwell K., Landy A.;
RT "Site-specific recombination functions of bacteriophage lambda: DNA
RT sequence of regulatory regions and overlapping structural genes for
RT int and xis.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:2482-2486(1980).
CC -I- FUNCTION: EXCISIONASE AND INTEGRASE ARE NECESSARY FOR THE EXCISION
CC OF PROPHAGE FROM THE HOST GENOME BY SITE-SPECIFIC RECOMBINATION AT
CC THE ATT SITE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J02459; AAA96563.1; -
CC DR PIR; A04321; RSBPXL.
CC KM DNA recombination; DNA-binding.
CC SQ SEQUENCE 72 AA; 8605 MW; 0E6A4843502200AA CRC64;

Query Match 35.4%; Score 46; DB 1; Length 72;
Best Local Similarity 47.8%; Pred. No. 2.6;
Matches 11; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 1 RRVVRRVRRVRRVRRVRRVRR 23
Db 6 QEWNAQR--PRLSTVRRVRR 26

RESULT 15
VIF_SIVAI
ID VIF_SIVAI STANDARD; PRT; 219 AA.
AC Q02841;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Vifion infectivity factor (SOR protein) (Q protein).
GN VIF.
OS Simian immunodeficiency virus (isolate AGM / clone GRI-1) (SIV-AGM).
OC Viruses; Retrovirus; Retroviridae; Lentiviridae.
OX NCBI_TaxID=31684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220680; PubMed=2024476;
RA Fomsgaard A., Hirsch V.M., Allan J.S., Johnson P.R.;
RT "A highly divergent proviral DNA clone of SIV from a distinct species
RT of African green monkey.";
RL Virology 182:397-402(1991).
CC -I- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC -I- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M66437; AAA91924.1; -
CC DR EMBL; M58410; AAA47589.1; -
CC DR InterPro; IPR000475; Viral_infect.
CC DR Pfam; PF00559; VIF; 1.
CC DR ProDom; PD000063; Viral_infect; 1.
CC KM AIDS.
CC SQ SEQUENCE 219 AA; 26087 MW; 20B3721B77EC9D4E CRC64;

Query Match 35.4%; Score 46; DB 1; Length 219;
Best Local Similarity 42.9%; Pred. No. 8.6;
Matches 9; Conservative 4; Mismatches 6; Indels 2; Gaps 1;

QY 1 RRVVRRVRRVRRVRRVRRVRR 21
Db 5 KQMT--VRVVRVRSERISRW 23
```

Search completed: June 9, 2003, 11:56:42
Job time : 6.85106 secs

```
CC VERNAMYCIN B, BUT NOT TO ALL MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN
CC B ANTIBIOTICS.
CC -1- INDUCTION: THE EXPRESSION OF CARB IS INDUCIBLE BY CERTAIN
CC MACROLIDE ANTIBIOTICS.
CC -1- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M16503; AAC32026.1; -.
CC PIR: A26512; A26512.
CC InterPro: IPR001737; RNA_A_dimeth.
CC InterPro: IPR000051; SAM_Bind.
CC Pfam: PF00398; RnaAD; 1.
CC PROSITE: PS01131; RNA_A_DIMETH; 1.
CC Antibioc resistance; Transferase; Methyltransferase; Plasmid.
SQ SEQUENCE 299 AA; 34138 MW; 63990A894C044C6 CRC64;

Query Match 36.5%; Score 47.5; DB 1; Length 299;
Best Local Similarity 39.3%; Pred. No. 7.4;
Matches 11; Conservative 1; Mismatches 11; Indels 5; Gaps 1;

QY 1 RRVRRRRRVRRV-----RVVRRRVRR 23
Db 183 RRVSRRLTVATVPEVEMWGERISRMRP 210

RESULT 12
MMJC_MYCTU STANDARD; PRT; 1146 AA.
AC OS0585;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative membrane protein mmpL12.
GN mmpL12 OR RV1522C OR WT1573 OR MTCY19G5.06.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Mycobacteriaceae; Mycobacterium.
OC Actinomycetales; Corynebacteriales; Mycobacteriales; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean A., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fietischmann R.D., Allard D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwim M.W., Halt D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Dichter A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
```

```
CC -1- SIMILARITY: BELONGS TO THE MPEL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z77826; CAB01394.1; -.
CC EMBL: AE007024; AAK45840.1; -.
CC TIGR: MT1573; -.
CC Tuberculosis; RV1522C; -.
CC InterPro: IPR004707; ActII.
CC InterPro: IPR000731; HMGCR/patch_5TM.
CC InterPro: IPR004859; MMEPL.
CC Pfam: PF03176; MPEL; 2.
CC TIGRPFAM: TIGR00833; actII; 1.
CC PROSITE: PS50156; SSD; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
FT TRANSMEM 254 274 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 382 402 POTENTIAL.
FT TRANSMEM 826 846 POTENTIAL.
FT TRANSMEM 850 870 POTENTIAL.
FT TRANSMEM 883 903 POTENTIAL.
FT TRANSMEM 928 948 POTENTIAL.
FT TRANSMEM 949 969 POTENTIAL.
SQ SEQUENCE 1146 AA; 122429 MW; AC03075641C5CB09 CRC64;

Query Match 36.5%; Score 47.5; DB 1; Length 1146;
Best Local Similarity 48.0%; Pred. No. 31;
Matches 12; Conservative 4; Mismatches 4; Indels 5; Gaps 2;

QY 1 RRVRRRRV-----RVVRRV-----RVVRR 20
Db 356 RGVKRPRLTSMRMRSGVRIYR 380

RESULT 13
VXIS_BP434 STANDARD; PRT; 72 AA.
ID VXIS_BP434
AC P11683; P16408;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Exci1onase.
GN XIS.
OS Bacteriophage 434, and
OS Bacteriophage HK022.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10712; 10742;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Phage 434;
RX MEDLINE=91346141; PubMed=1715186;
RA Baker J., Limberger R., Schneider S.J., Campbell A.;
RT "Recombination and modular exchange in the genesis of new lambdoid
RT phages."
RL New Biol. 3:297-308(1991).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC SPECIES=Phage 434;
RX MEDLINE=88167849; PubMed=2965063;
RA Limberger R.J., Campbell A.M.;
RT "Functional elements of DNA upstream from the integrase operon that
RT are conserved in bacteriophages 434 and lambda."
RL Gene 61:135-144(1987).
RN [3]
```

16-OCT-2001 (Rel. 40, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Myosin XV (Unconventional myosin-15).
 GN MYO15A OR MYO15.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=20021762; PubMed=10552926;
 RA Liang Y., Wang A., Belyantseva I.A., Anderson D.W., Probst F.J.,
 RA Barber J.D., Miller W., Touchman J.W., Jin L., Sullivan S.L.,
 RA Seiders J.R., Camper S.A., Lloyd R.V., Kachar B., Friedman T.B.,
 RA Fridell R.A.;
 RT "Characterization of the human and mouse unconventional myosin XV
 RT genes responsible for hereditary deafness DFNB3 and shaker 2.",
 RL Genomics 61:243-258(1999).
 [2]
 RP SEQUENCE OF 1168-2970 FROM N.A., AND VARIANT SH2 TYR-1779.
 RC TISSUE=Embryo;
 RX MEDLINE=98267310; PubMed=9603735;
 RA Probst F.J., Fridell R.A., Raphael Y., Saunders T.L., Wang A.,
 RA Liang Y., Morell R.J., Touchman J.W., Lyons R.H., Noden-Flaugh K.,
 RA Friedman T.B., Camper S.A.;
 RT "Correction of deafness in shaker-2 mice by an unconventional myosin
 RT in a BAC transgene.",
 RL Science 280:1444-1447(1998).
 [3]
 RP SEQUENCE OF 1237-1823 FROM N.A., AND VARIANT SH2 TYR-1779.
 RC STRAIN=C57BL/6;
 RX MEDLINE=98369604; PubMed=9703981;
 RA Wakabayashi Y., Takahashi Y., Kikawa Y., Okano H., Mishima Y.,
 RA Ushiki T., Yonekawa H., Komitani R.;
 RT "A novel type of myosin encoded by the mouse deafness gene shaker-2.",
 RL Biochem. Biophys. Res. Commun. 248:655-659(1998).
 CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
 CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
 CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
 CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
 CC SIMILARITY). MAY PLAY A ROLE IN THE FORMATION OR MAINTENANCE OF
 CC THE ACTIN-RICH STRUCTURES OF THE INNER EAR SENSORY HAIR CELLS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: IN THE DEVELOPING INNER EAR, EXPRESSED IN
 CC COCHLEA AND VESTIBULAR APPARATUS. EXPRESSION APPEARS TO BE
 CC RESTRICTED TO COCHLEAR NEUROSENSORY CELLS AND UPPER EPITHELIAL
 CC LAYER OF MACULA SACCUA. ALSO EXPRESSED IN VACUOLA UTICULIT AND
 CC CRISTAE AMPULLARIS OF THE SEMICIRCULAR CANALS. IN ADULT COCHLEAR
 CC HAIR CELLS, HIGHEST EXPRESSION IN STEREOCILIA AND APICAL BODY.
 CC -1- DISEASE: DEFECTS IN MYO15A ARE THE CAUSE OF SHAKER 2 (SH2), A
 CC CONDITION CAUSING DEAFNESS, CIRCLING BEHAVIOR, HEADDOSSING AND
 CC HYPERACTIVITY. AUDITORY HAIR CELLS OF AFFECTED ANIMALS HAVE VERY
 CC SHORT STEREOCILIA AND A LONG ACTIN-CONTAINING PROTRUSION AT THEIR
 CC BASAL END.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IO DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC EMBL; AF144095; AAF05904.1; -
 CC EMBL; AF053130; AAC40124.1; -
 CC EMBL; AB014510; BAB36582.1; -
 CC HSSP; P10587; 1BR2.
 CC MGD; MG1:1261811; Myo15.

InterPro: IPR000299; Band_4.1.
 DR InterPro: IPR000048; IO_region.
 DR InterPro: IPR000857; MYTH4.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR001609; myosin head.
 DR Pfam: PF00063; myosin_head_1.
 DR Pfam: PF00612; IO_2.
 DR Pfam: PF00784; MYTH4_2.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO_2.
 DR SMART: SM00242; MYSC; 1.
 DR SMART: SM00139; MYTH4_2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50096; IO_2.
 DR PROSITE: PS50002; SH3; 1.
 DR PROSITE: PS00660; BAND_4.1; FALSE_NEG.
 DR PROSITE: PS00661; BAND_41_2; FALSE_NEG.
 DR PROSITE: PS50057; BAND_41_3; FALSE_NEG.
 KW Myosin; ATP-binding; Actin-binding; Coiled coil; Repeat; SH3 domain;
 KW Disease mutation; Deafness.
 FT DOMAIN 1 1871
 FT DOMAIN 1872 2013
 FT DOMAIN 2014 3511
 FT DOMAIN 1307 1334
 FT DOMAIN 1776 1783
 FT DOMAIN 1886 1908
 FT DOMAIN 1909 1938
 FT DOMAIN 2848 2934
 FT DOMAIN 3187 3424
 FT NP_BIND 1299 1306
 FT VARIANT 1779 1779
 FT CONFLICT 1330 1331
 FT CONFLICT 1579 1579
 FT CONFLICT 1955 1972
 FT CONFLICT 2077 2077
 FT CONFLICT 2139 2139
 FT CONFLICT 2953 2953
 SQ SEQUENCE 3511 AA; 395533 MW; 38C962F98A2D395B CRC64;
 Query Match 36.9%; Score 48; DB 1; Length 3511;
 Best Local Similarity 43.5%; Pred. No. 87;
 Matches 10; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
 Qy 6 RVRRVRRV-----RVRRVRRV 24
 Db 858 RLPRTRRLSEPTPLAVKRVRR 880
 RESULT 11.
 CARB_STRTH STANDARD; PRT; 299 AA.
 ID P13079;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE rRNA methyltransferase (EC 2.1.1.-) (Carbomycin-resistance protein).
 GN CARB.
 OS Streptomyces thermotolerans.
 OC Plasmid pOU159.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=80858;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11416;
 RX MEDLINE=8724811; PubMed=3036668;
 RA Epp J.K., Burgett S.G., Schoner B.E.;
 RT "Cloning and nucleotide sequence of a carbomycin-resistance gene from
 RT Streptomyces thermotolerans.",
 RL Gene 53:73-83(1987).
 CC -1- FUNCTION: PROBABLE RNA METHYLASE. CARB CONFERS RESISTANCE TO
 CC CARBOMYCIN AND SEVERAL OTHER MACROLIDES, LINCOMYCIN AND

KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 105 AA; 12527 MW; 7E5DF7999E74A098 CRC64;
Query Match 41.2%; Score 53.5; DB 1; Length 105;
Best Local Similarity 59.1%; Pred. No. 0.35;
Matches 13; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
QY 3 WVRVRVRVRVRVRVRVRVR 24
DB 19 WGRRTTTRRA-IRAVRVVREFFVR 39
RESULT 6
ID Y278_MYCTU STANDARD; PRT; 957 AA.
AC P56877;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Hypothetical PE-PGRS family protein RV0278c precursor.
GN RV0278C OR MT0291 OR MTV035_06C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae;
OC Actinomycetidae; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Bishmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Baaham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gencles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagsels K., Kitchin A., McLean J., Moul S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Batteil B.G.;
RT "Deciphering the Biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AL021930; CA11735.1; -
CC EMBL; AE006936; AAK44511.1; ALT_INIT.
CC TIGR; MT0291;
DR Tuberculosis; RV0278c; -
DR InterPro; IPR000084; PE_region.
DR Pfam; PF00934; PE_1.
KW Hypothetical protein; Repeat; Signal; Complete proteome.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 957 HYPOTHETICAL PE-PGRS FAMILY PROTEIN
FT RV0278C.
FT CONFLICT 40 40 M-> I (IN REF. 2).

FT CONFLICT 158 163 MISSING (IN REF. 2).
FT CONFLICT 807 807 R-> G (IN REF. 2).
SQ SEQUENCE 957 AA; 81905 MW; 71EBABD417FBA47C CRC64;
Query Match 39.2%; Score 51; DB 1; Length 957;
Best Local Similarity 64.7%; Pred. No. 8.3;
Matches 11; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
QY 8 RRVVRVRVRVRVRVRVR 24
DB 881 RRVVRVRVR--QRMCR 895
RESULT 7
ID YU30_RALSO STANDARD; PRT; 1582 AA.
AC O8XV02;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Hypothetical protein RSC3030 precursor.
GN RSC3030 OR RS04727.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brothier P., Camus J.C., Cattellico L.,
RA Chandler M., Choigne N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspard C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Signier P., Thibault P., Whalen M., Wincker P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -1- SIMILARITY: BELONGS TO THE UPF0192 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AL646073; CAD16739.1; -
CC EMBL; AF000000; Signal; Complete proteome.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 1582 HYPOTHETICAL PROTEIN RSC3030.
SQ SEQUENCE 1582 AA; 170090 MW; 8683D146F5A3CA CRC64;
Query Match 39.2%; Score 51; DB 1; Length 1582;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 3 WVRVRVRVRVRVRVRVR 20
DB 125 WTRORARVORVLSFATR 142
RESULT 8
ID HKXL_ARATH STANDARD; PRT; 493 AA.
AC O9T071;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable hexokinase (EC 2.7.1.1).
GN AT4G37840 OR T28119.120.
OS Arabidopsis thaliana (Mouse-ear cress).

RX MEDLINE=97207018; PubMed=9054511;
 RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
 RT "Insights into multisite phosphorylation from the crystal structure of
 the C-terminal HPC domain of ArcB.";
 RL Cell 88:717-723(1997).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 659-776 IN COMPLEX WITH CHEX.
 RX MEDLINE=98437504; PubMed=9761838;
 RA Kato M., Mizuno T., Hakoshima T.;
 RT "Crystallization of a complex between a novel C-terminal transmitter,
 HPC domain, of the anaerobic sensor kinase ArcB and the chemotaxis
 response regulator CheY.";
 RL Acta Crystallogr. D 54:140-142(1998).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 659-776.
 RX MEDLINE=20003135; PubMed=10531481;
 RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
 RT "Refined structure of the histidine-containing-phosphotransfer (HPC)
 domain of the anaerobic sensor kinase ArcB from Escherichia coli at
 1.57-A resolution.";
 RL Acta Crystallogr. D 55:1842-1849(1999).
 CC -1- FUNCTION: Member of the two-component regulatory system arcB/arcA.
 CC Sensor-regulator protein for anaerobic repression of the arc
 CC modulation. Activates arcA via a four-step phosphorylation. ArcB can
 CC also dephosphorylate arcA by a reverse phosphorylation involving His-
 CC 717 and Asp-576.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Probable).
 CC -1- PTM: Activation requires a sequential transfer of a phosphate
 CC group from a His in the primary transmitter domain, to a Asp in
 CC the receiver domain and to a His in the secondary transmitter
 CC domain.
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X53315; CA37397.1; -;
 DR EMBL; U18997; AA58012.1; -;
 DR EMBL; AE000400; AAC76242.1; -;
 DR PIR; S11794; RGEAR.
 DR PDB; 1A0B; 18-MAR-98.
 DR PDB; 2A0B; 17-JUN-98.
 DR PDB; 1BDY; 11-MAY-99.
 DR ECGene; EGI0062; arcB.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR004359; HIS_KIN_slg.
 DR InterPro; IPR003661; His_KinA.
 DR InterPro; IPR002570; Hpt.
 DR InterPro; IPR000700; PAS-assoC.
 DR InterPro; IPR000014; PAS_domain.
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF00072; response_reg; 2.
 DR Pfam; PF00512; signal; 1.
 DR Pfam; PF02518; HATPase_C; 1.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00387; HATPase_C; 1.
 DR SMART; SM00073; HPT; 2.
 DR SMART; SM00388; HisKA; 1.
 DR SMART; SM00091; PAS; 1.
 DR SMART; SM00448; REC; 2.
 DR TIGRfam; TIGR00223; sensory_box; 1.
 DR PROSITE; PS0109; HIS_KIN; 1.
 DR PROSITE; PS0113; PAC; 1.

DR PROSITE; PS0112; PAS; 1.
 DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
 KW Sensory transduction; Transferase; Kinase; Phosphorylation;
 KW Transmembrane; Inner membrane; Transcription regulation; 3D-structure;
 KW Complete proteome.
 KM
 FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 26 46 POTENTIAL.
 FT DOMAIN 47 57 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 58 78 POTENTIAL.
 FT DOMAIN 79 78 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 153 223 PAS.
 FT DOMAIN 226 278 PAC.
 FT DOMAIN 289 507 HISTIDINE KINASE.
 FT DOMAIN 527 643 SECONDARY TRANSMITTER DOMAIN (POTENTIAL).
 FT DOMAIN 644 778 PHOSPHORYLATION (AUTO-).
 FT MOD RES 292 292 PHOSPHORYLATION (PROBABLE).
 FT MOD RES 576 576 PHOSPHORYLATION (PROBABLE).
 FT MOD RES 717 717 H->Q: LOSS OF ACTIVITY.
 FT MUTAGEN 292 292 H->A: LOSS OF ACTIVITY.
 FT MUTAGEN 576 576 H->Q: LOSS OF ACTIVITY.
 FT MUTAGEN 717 717 MISSING (IN REF. 2).
 FT CONFLICT 459 470
 SQ SEQUENCE 778 AA; 87982 MW; DD6LEAECE95AD30 CR664;
 Query Match 41.5%; Score 54; DB 1; Length 778;
 Best Local Similarity 31.8%; Pred. No. 2.5;
 Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
 QY 3 WRRRVRRVRRVRRVRRVRR 24
 DB 753 WIEEMKEWRHDEVLKAWAK 774
 ID RL31_AERPE STANDARD; PRT; 105 AA.
 AC Q9YD25;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 50S ribosomal protein L31e.
 GN RPL31E OR APE1087.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
 OC Desulfurococcaceae; Aeropyrum.
 OC NCBI_TaxID=56636;
 OK
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=X1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y.,
 RA Jinno K., Takahashi M., Sekine M., Baba S.-T., Anai A., Kousugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix X1.";
 RL DNA Res. 6:83-101(1999).
 CC -1- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AP000060; BAA8072.1; -;
 DR InterPro; IPR000054; Ribosomal_L31e.
 DR Pfam; PF01198; Ribosomal_L31e; 1.
 DR PROSITE; PS01144; RIBOSOMAL_L31E; FALSE_NEG.

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blatner F.R.,
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shingawa H.,
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: Member of the two-component regulatory system arcB/arcA.
 CC Sensor-regulator protein for anaerobic repression of the arc
 CC modulon. Activates arcA via a four-step phosphoryl. ArcB can
 CC also dephosphorylate arcA by a reverse phosphoryl. involving His-
 CC 717 and Asp-576 (by similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Probable).
 CC -1- PTM: Activation requires a sequential transfer of a phosphate
 CC group from a His in the primary transmitter domain, to a Asp in
 CC the receiver domain and to a His in the secondary transmitter
 CC domain (by similarity).
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AE005549; BAB37512.1; -;
 DR EMBL; AB002564; BAB37512.1; -;
 DR InterPro: IPR003594; ATPrind_ATPase.
 DR InterPro: IPR004359; HIS_KIN_519.
 DR InterPro: IPR003661; His_KinA.
 DR InterPro: IPR002570; Hpt.
 DR InterPro: IPR000700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS domain.
 DR InterPro: IPR001789; Response_reg.
 DR Pfam; PF00072; response_reg; 1.
 DR Pfam; PF00512; signal; 1.
 DR Pfam; PF00989; PAS; 1.
 DR Pfam; PF02518; HATPase_C; 1.
 DR SMART; SM00387; HATPase_C; 1.
 DR SMART; SM00073; HPT; 1.
 DR SMART; SM00388; HisKA; 1.
 DR SMART; SM00091; PAS; 1.
 DR TIGRfam; TIGR00229; sensory_box; 1.
 DR PROSITE; PS50109; HIS_KIN_1.
 DR PROSITE; PS50113; PAC; 1.
 DR PROSITE; PS50112; PAS; 1.
 DR PROSITE; PS50110; RESPONSE REGULATORY; 1.
 DR Sensory transduction; Transferrase; Kinase; Phosphorylation;
 KW Transmembrane; Inner membrane; Transcription regulation;
 KW Complete proteome.
 KM
 FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 26 46 POTENTIAL.
 FT DOMAIN 47 57 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 58 78 POTENTIAL.
 FT DOMAIN 79 778 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 153 223 PAS.
 FT DOMAIN 226 278 PAC.
 FT DOMAIN 289 507 HISTIDINE KINASE.

FT DOMAIN 527 643 RESPONSE REGULATORY.
 FT DOMAIN 644 778 SECONDARY TRANSMITTER DOMAIN (POTENTIAL).
 FT MOD RES 292 292 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD RES 576 576 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD RES 717 717 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 778 AA; 88010 MW; C8AE004B007F930 CRC64;
 Query Match 41.5%; Score 54; DB 1; Length 778;
 Best local Similarity 31.8%; Pred. No. 2.5;
 Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
 QY 3 WVRVRVRVRVRVRVRVRVR 24
 DB 753 WIEEMKEWRHVEVLKAVVAK 774
 ID ARCB_ECOLI STANDARD; PRT; 778 AA.
 AC P22763;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aerobic respiration control sensor protein arcB (EC 2.7.3.-).
 GN ARCB OR B3210.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=90355832; PubMed=2201868;
 RA Iuchi S., Matsuda Z., Fujiwara T., Lin E.C.C.;
 RT "The arcB gene of Escherichia coli encodes a sensor-regulator protein
 RT for anaerobic repression of the arc modulon";
 RL Mol. Microbiol. 4:715-727(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP CHARACTERIZATION.
 RC STRAIN=M15;
 RX MEDLINE=97431492; PubMed=9286997;
 RA Georgellis D., Lynch A.S., Lin E.C.C.;
 RT "In vitro phosphorylation study of the arc two-component signal
 RT transduction system of Escherichia coli";
 RL J. Bacteriol. 179:5429-5435(1997).
 RN [4]
 RP CHARACTERIZATION.
 RC STRAIN=M15;
 RX MEDLINE=99047671; PubMed=9830034;
 RA Georgellis D., Kwon O., De Wulf P., Lin E.C.C.;
 RT "Signal decay through a reverse phosphoryl. in the arc two-component
 RT signal transduction system";
 RL J. Biol. Chem. 273:32864-32869(1998).
 RN [5]
 RP MUTAGENESIS OF HIS-292; ASP-576 AND HIS-717.
 RC STRAIN=K12 / MC4100;
 RX MEDLINE=20309722; PubMed=10851007;
 RA Kwon O., Georgellis D., Lin E.C.C.;
 RT "Phosphorelay as the sole physiological route of signal transduction
 RT by the arc two-component system of Escherichia coli";
 RL J. Bacteriol. 182:3858-3862(2000).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.06 ANGSTROMS) OF 660-778.

```

RN [2] IDENTIFICATION, AND TISSUE SPECIFICITY.
RP MEDLINE=20175760; PubMed=10710312;
RX Clyne P.J., Warr C.G., Carlson J.R.;
RA "Candidate taste receptors in Drosophila.";
RL Science 287:1830-1834(2000).
RN [3]
RP IDENTIFICATION.
RX MEDLINE=21407712; PubMed=11516643;
RA Dunipace L., Meister S., McNeely C., Amrein H.;
RT "Spatially restricted expression of candidate taste receptors in the
RL Drosophila gustatory system.";
RN Curr. Biol. 11:822-835(2001).
RN [4]
RP CONCEPTUAL TRANSLATION.
RA Robertson H.;
RL unpublished observations (NOV-2001).
CC -1- FUNCTION: Probable role in the gustatory response.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- TISSUE SPECIFICITY: Expressed in the adult labellar chemosensory
CC neurons.
CC -1- SIMILARITY: BELONGS TO FAMILY DR-TR OF G-PROTEIN COUPLED
CC RECEPTORS. SUBFAMILY VI.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE003459; AAF46958.2; ALT_SEQ.
DR FlyBase; FBgm004123; Gr59c.
KM Hypothetical protein; Receptor; G-protein coupled receptor;
KM Transmembrane; Glycoprotein; Multigene family.
FT DOMAIN 1 3
FT TRANSSEM 4 24
FT DOMAIN 25 39
FT TRANSSEM 40 60
FT DOMAIN 61 75
FT TRANSSEM 76 96
FT DOMAIN 97 166
FT TRANSSEM 167 187
FT DOMAIN 188 259
FT TRANSSEM 260 280
FT DOMAIN 281 284
FT TRANSSEM 285 305
FT DOMAIN 306 372
FT TRANSSEM 373 393
FT DOMAIN 394 397
FT CARBOHD 61 61
SQ SEQUENCE 397 AA; 46164 MW; 60CEFC8AC46FD0E CRC64;

Query Match 43.8%; Score 57; DB 1; Length 397;
Best local Similarity 36.4%; Pred. No. 0.477;
Matches 12; Conservative 6; Mismatches 5; Indels 10; Gaps 2.

Cy 2 RWVRVR--RWVRVRVVRVVR-----RWVR 24
||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 97 RWQGRSRFRIRIMQIILVVRDPPQVVRGRWVR 129

RESULT 2
ID SECE_SUTSO STANDARD; PRT; 62 AA.
AC P58131;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Preproctein, translocase sec6 subunit (protein transport protein SEC61

```

DE gamma subunit homolog) .
 GN SECE OR SSO5663.
 OS Sulfolobus solfataricus.
 CC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 CC Sulfolobus.
 CC NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=1142776;
 RA Shew Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Mayez W.J., Chan-Welner C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erusso G., Fletcher A.C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA The-Noc H.P., Redder P., Schenk M.E., Thériault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gassnerland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -1- FUNCTION: Involved in protein export (By similarity).
 CC -1- SUBUNIT: Component of the protein translocase complex (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Tail-anchored membrane protein (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE SECE/SECE1-GAMMA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; AE006668; AAK40677.1; -.
 DR InterPro; IPR0041901; SECE.
 DR InterPro; IPR004795; SECE_euk_arch.
 DR TIGRFAMs; TIGR00327; sece_euk_arch.1.
 DR PROSITE; PS01067; SECE_SECE1G; FALSE_NEG.
 KW Protein translocator; Translocation; Transmembrane; Complete proteome.
 FT TRANSMEM 40 60
 FT POTENTIAL.
 SQ SEQUENCE 62 AA; 7184 MW; CDS1874F424BE9FA CRC64;

 OY Query Match 42.3%; Score 55; DB 1; Length 62;
 Best Local Similarity 44.4%; Pred. No. 0.12;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

 Db 3 WVRVRVRRVRRVVRVVR 20
 9 WFRRLREDNRRIITVARK 26

 RESULT 3
 ARCB_ECO57
 ID ARCB_ECO57 STANDARD; PRT; 778 AA.
 AC P58363;
 DT 15-JUN-2002 (Rel. 41, Created).
 DT 15-JUN-2002 (Rel. 41, Last sequence update).
 DT 15-JUN-2002 (Rel. 41, Last annotation update).
 DE Aerobic respiration control sensor protein arcb (EC 2.7.3.-).
 GN ARCB OR 24574 OR ECS4089.
 OS Escherichia coli O157:H7.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Pena N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobbeck E.T., Davis N.W., Lim A., Dimilanta E.T., Potamoustis K.,

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:40:31 / Search time 4.85106 Seconds
(without alignments)
205.199 Million cell updates/sec

Title: US-10-079-075-10

Perfect score: 130
Sequence: 1 RRMVRRVRRVRRVRRVRRVRRVRR 24

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	43.8	397	1 G59C_DROME	Q9W1U5 drosophila
2	55	42.3	62	1 SECE_SUISO	P88191 sulfolobus
3	54	41.5	778	1 ARCB_ECO57	P83633 escherichia
4	54	41.5	778	1 ARCB_ECOLI	P22763 escherichia
5	53.5	41.2	105	1 RL31_AERPE	Q9Y245 aeropyrum p
6	51	39.2	957	1 Y278_MYCTU	P56877 mycobacteri
7	51	39.2	1582	1 YU30_RALSO	Q8XV02 ralsonia s
8	49	37.7	493	1 HXKL_ARATH	Q9T071 arabidopsis
9	48.5	37.3	213	1 URK_MYCPN	P75217 mycoplasma
10	48	36.9	3511	1 MY15_MOUSE	Q9QZ24 mus muscucu
11	47.5	36.5	299	1 CARB_STRTH	P13079 streptomyce
12	47.5	36.5	1146	1 MMLC_MYCTU	O50585 mycobacteri
13	46	35.4	72	1 VXIS_BP334	P11683 bacterioph
14	46	35.4	72	1 VXIS_LAMBD	P33659 bacterioph
15	46	35.4	219	1 VIF_STVAV	Q02841 simian immu
16	46	35.4	470	1 YUIR_ECOLI	P02841 escherichia
17	45.5	35.0	213	1 URK_MYCPN	P47652 mycoplasma
18	45	34.6	160	1 GPH_PSESS	P42510 pseudomonas
19	45	34.6	958	1 MMLI_MYCTU	P95211 mycobacteri
20	44.5	34.2	151	1 RL19_SUISO	Q9UX99 sulfolobus
21	44.5	34.2	155	1 RL19_AERPE	Q9Y193 aeropyrum p
22	44.5	34.2	523	1 NCAP_MEASE	P04851 measles vir
23	44.5	34.2	525	1 NCAP_MEASH	P25972 measles vir
24	44.5	34.2	525	1 NCAP_MEASH	P10050 measles vir
25	44.5	34.2	525	1 NCAP_MEASH	P26029 measles vir
26	44.5	34.2	525	1 NCAP_MEASH	P26030 measles vir
27	44.5	34.2	525	1 NCAP_MEASH	P26030 measles vir
28	44	33.8	65	1 SECE_SUJAC	P03354 rous sarcom
29	44	33.8	357	1 RPB_PASMU	P07340 sulfolobus
30	44	33.8	384	1 RPD_KLITA	Q9C986 pasteurella
31	44	33.8	452	1 PPOK_MYCTU	P41771 kluyteromyc
32	44	33.8	477	1 Y006_CAEEL	O53230 mycobacteri
33	44	33.8	556	1 PDPK_HUMAN	Q09236 caenorhabdi
					O15530 homo sapien

34	44	33.8	559	1 PDPK_MOUSE	Q92240 mus muscucu
35	44	33.8	559	1 PDPK_RAT	O55173 rattus norv
36	44	33.8	628	1 SYI_NOSLO	O27707 nosema locu
37	44	33.8	967	1 MML4_MYCTU	O53735 mycobacteri
38	43.5	33.5	254	1 GTXC_ORYSA	O06398 oryza sativ
39	43.5	33.5	357	1 LIPB_MYXMA	O9X654 myxococcus
40	43.5	33.5	745	1 NFRB_ECOLI	P31599 escherichia
41	43	33.1	33	1 PRTB_MUGCE	P08130 mugil cepha
42	43	33.1	85	1 Y024_BPTA	P39231 bacterioph
43	43	33.1	187	1 RS4_METUA	P54020 methanococc
44	43	33.1	243	1 FOL3_HUMAN	P41439 homo sapien
45	43	33.1	272	1 GPHI_PSEAR	Q96586 pseudomonas

ALIGNMENTS

RESULT 1
ID G59C_DROME STANDARD; PRT; 397 AA.
AC Q9W1U5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Putative gustatory receptor 59c.
GN G59C OR GR59D.2 OR CG13543 OR CG30186.
OS Drosophila melanogaster (Fruit fly).
OC Insecta; Pterygota; Neoptera; Endopterygota; Pancrustacea; Hexapoda;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutcliffe G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ayll J.F., Agdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fester C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kemmison J.A., Ketchum K.A.,
RA Kamali B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostreli A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacible J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Styrkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).

OY 3 WVRVRVRVRVRVRVRVR 21
 ||:|:|:|:|:|:|
 Db 753 WDELKLEKRNNDQVLRW 771

RESULT 12

hypothetical glycine-rich protein RV0278c - Mycobacterium tuberculosis (strain H37RV)
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C/Accession: D70835
 R/Collection: S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Squares, R., Sulton, J.E., Taylor, K., Whitehead, S., Barrett, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500, PMID:98295987, PMID:9634230
 A/Accession: D70835
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1957 <COL>
 A/Cross-references: GB:AL021930; GB:AL123456; NID:g3261524; PIDN:CA117353.1; PID:g290946
 A/Experimental source: strain H37RV
 C/Genetics:
 A/Gene: RV0278c
 C/Superfamily: elastin

Query Match 39.2%; Score 51; DB 2; Length 957;
 Best Local Similarity 64.7%; Pred. No. 41;
 Matches 11; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

OY 8 RRVRRVRVRVRVRVRVR 24
 |||||:|:|:|:|:|:|
 Db 881 RRVRRVRVR--QKRCR 895

RESULT 13

A59251
 myosin - Acetabularia cliftonii
 C/Species: Acetabularia cliftonii
 C/Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000
 C/Accession: A59251
 R/Vigrek, O.; Menzel, D.
 submitted to Genbank, March 1997
 A/Description: Molecular analysis of the cDNA coding for an unconventional myosin from
 A/Reference number: A59251
 A/Accession: A59251
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-1145 <YUG>
 A/Cross-references: GB:U94398; NID:g2051982; PIDN:AA53062.1; PID:g2051983
 C/Genetics:
 A/Gene: myo2
 A/Gene: myo2
 A/Gene: myo2
 A/Genetic code: SGC5
 C/Superfamily: myosin motor domain homology
 F;97-799/Domain: myosin motor domain homology <WMO>

Query Match 38.5%; Score 50; DB 2; Length 1145;
 Best Local Similarity 40.7%; Pred. No. 64;
 Matches 11; Conservative 7; Mismatches 5; Indels 4; Gaps 2;

OY 1 RRVRRVR--RRVRRVR--RVRRVR 23
 ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
 Db 914 QKRRFRVQSRNRNKKVIAAIIQKFR 940

RESULT 14

T06031
 hexokinase homolog T28119.120 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999
 C/Accession: T06031

R/Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voel, M.; Robben, J.; Volckaert, G.; Ba
 submitted to the Protein Sequence Database, March 1999
 A/Reference number: Z15484
 A/Accession: T06031

A/Molecule type: DNA
 A/Residues: 1-493 <BEV>
 A/Cross-references: EMBL:AL035709; GSPDB:GN00062; ATSP:T28119.120
 A/Experimental source: cultivar Columbia; BAC clone T28119
 C/Genetics:
 A/Gene: ATSP:T28119.120
 A/Map position: 4
 A/Intons: 92/2; 142/3; 227/3; 279/3; 312/3; 356/2; 383/3
 C/Superfamily: hexokinase, hexokinase homology

Query Match 37.7%; Score 49; DB 2; Length 493;
 Best Local Similarity 36.4%; Pred. No. 41;
 Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 2 RRVRRVRVRVRVRVRVRVR 23
 ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
 Db 26 RVRRRRRLKHTQRLRKFR 47

RESULT 15

T08179
 LRGS protein - Chlamydomonas reinhardtii
 C/Species: Chlamydomonas reinhardtii
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
 C/Accession: T08179
 R/Gloeckner, G.; Beck, C.F.
 submitted to the EMBL Data Library, October 1996
 A/Description: Molecular characterization of a gene (LRGS) involved in blue light signal
 A/Reference number: Z16399
 A/Accession: T08179
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-640 <GLO>
 A/Cross-references: EMBL:U73817; NID:g1644369; PID:g1644370
 C/Genetics:
 A/Gene: LRGS

Query Match 37.7%; Score 49; DB 2; Length 640;
 Best Local Similarity 60.0%; Pred. No. 52;
 Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 5 RRVRRVRVRVRVRVRVRVR 24
 ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
 Db 460 RRLQRRRRRGGRVRVR 479

Search completed: June 9, 2003, 12:03:15
 Job time: 10.9574 secs

RESULT 7
H72708
probable ribosomal protein L31 APE1087 - Aeropyrum pernix (strain K1)
C/Species: Aeropyrum pernix
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 02-Aug-2002
C/Accession: H72708
R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Takenawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawada, Res. 6, 83-101, 1999
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A/Reference number: A72450; PMID:99310339; PMID:10382966
A/Accession: H72708
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-105 <RAW>
A/Cross-references: DDBJ:AF000060; NID:95104188; PIDN:BAA80072.1; PID:d1043858; PID:95104188
A/Experimental source: strain K1
C/Genetics:
A/Gene: APE1087
C/Superfamily: rat ribosomal protein L31

Query Match 41.2%; Score 53.5; DB 2; Length 105;
Best Local Similarity 59.1%; Pred. No. 2.7;
Matches 13; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 3 WVRVRVRVRVRVRVRVRVR 24
DB 19 WGRRTTAA-IRAVRMREFVR 39

RESULT 8
H70846
hypothetical glycine-rich protein RV3345c - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C/Accession: H70846
R/Cole, S.T.; Broesch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Sulton, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; PMID:9825987; PMID:9634230
A/Accession: H70846
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-1538 <CON>
A/Cross-references: GB:AL021841; GB:AL123456; NID:93261517; PIDN:CA17117.1; PID:93261517
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: RV3345c
C/Superfamily: collagen alpha 1(IV) chain

Query Match 41.2%; Score 53.5; DB 2; Length 1538;
Best Local Similarity 46.2%; Pred. No. 29;
Matches 12; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

QY 2 RW---VRRVRVRVRVRVRVRVR 24
DB 1488 RWRCRRLKRWKQRCRRCPWLR 1513

RESULT 9
A34413
atrial gland granule-specific antigen precursor - California sea hare
C/Species: Aplysia californica (California sea hare)
C/Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 21-Jul-2000
C/Accession: A34413
R/Sossein, W.S.; Kreiner, T.; Barnaga, M.; Schilling, J.; Scheller, R.H.
J. Biol. Chem. 264, 16933-16940, 1989
A/Title: A dense core vesicle protein is restricted to the cortex of granules in the exo
A/Reference number: A34413; PMID:89380331; PMID:2777814

A/Accession: A34413
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-432 <SOS>
A/Cross-references: GB:J05059; NID:9155709; PIDN:AAA27741.1; PID:9155710
C/Superfamily: California sea hare atrial gland granule-specific antigen

Query Match 39.2%; Score 51; DB 2; Length 432;
Best Local Similarity 26.3%; Pred. No. 20;
Matches 5; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 3 WVRVRVRVRVRVRVRVRVR 21
DB 401 WPKWESLWRSILSLRSW 419

RESULT 10
T32743
hypothetical protein F57B10.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Jun-2000
C/Accession: T32743
R/Greco, T.; Elliott, G.; Keppeler, D.
submitted to the EMBL Data Library, December 1997
A/Description: The sequence of C. elegans cosmid F57B10.
A/Reference number: Z21219
A/Accession: T32743
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-570 <GRE>
A/Cross-references: EMBL:AF039713; PIDN:AMB96724.1; GSPDB:GN00019; CESP:F57B10.7
A/Experimental source: strain Bristol N2; clone F57B10
C/Genetics:
A/Gene: CESP:F57B10.7
A/Map position: 1
A/Insertion: 10/2; 44/1; 118/3; 152/2; 183/2; 324/2; 425/3; 522/3
C/Superfamily: human alpha, alpha-trehalase

Query Match 39.2%; Score 51; DB 2; Length 570;
Best Local Similarity 42.1%; Pred. No. 26;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 RWRVRVRVRVRVRVRVRVR 19
DB 102 RRMALHLRIMWDLCKRV 120

RESULT 11
AD0432
aerobic respiration control sensor/response regulatory protein (EC 2.7.3.-) [imported]
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C/Accession: AD0432
R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibbali, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; 11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; PMID:21470413; PMID:11586360
A/Cross-references: GB:AL590842; PIDN:CA92784.1; PID:915981477; GSPDB:GN00175
A/Accession: AD0432
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-778 <KUR>
A/Cross-references: GB:AL590842; PIDN:CA92784.1; PID:915981477; GSPDB:GN00175
C/Genetics:
A/Gene: arcB
C/Superfamily: aerobic respiration control sensor protein arcB; response regulator hmoJ
C/Keywords: phosphotransferase

Query Match 39.2%; Score 51; DB 2; Length 778;
Best Local Similarity 42.1%; Pred. No. 34;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

A:Experimental source: strain Danver Half-long

Query Match

Best Local Similarity 44.2%; Score 57.5; DB 2; Length 111;

Matches 14; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 2 RWVRVRVRV-RRVVRVRVRVR 23

DB 65 RMRRRRRVRRRRRRRRRRRCR 87

RESULT 3

P90177 hypothetical protein secc [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 17-May-2002

C/Accession: P90177

R/Shu, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.U.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.

arrest, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

Submitted to Genbank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A91319

A:Accession: P90177

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-62 <KUR>

A:Cross-references: GB:AE006641; NID:g13813489; PIDN:AAK40677.1; GSPDB:GN00155

C:Genetics:

A:Gene: secc

C:Superfamily: yeast SSI1 protein

Query Match

Best Local Similarity 42.3%; Score 55; DB 2; Length 62;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 WVRVRVRVRVRVRVRVR 20

DB 9 WFRRLRDMNRITVARK 26

RESULT 4

RGECAR

aerobic respiration control sensor protein arcB (EC 2.7.3.-) - Escherichia coli (strain

C:Species: Escherichia coli

C>Date: 31-Dec-1991 #sequence_revision 17-Oct-1997 #text_change 01-Mar-2002

C/Accession: D65112; J00295; S11799

R/Baltcher, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: D65112

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-776 <BAT>

A:Cross-references: GB:AE004000; GB:U00096; NID:92367203; PIDN:AACT6242.1; PID:g1789603;

A:Experimental source: strain K-12, substrain MG1655

R/Inchi, S.; Matsuda, Z.; Fujiwara, T.; Lin, E.C.C.

Mol. Microbiol. 4, 715-727, 1990

A:Title: The arcB gene of Escherichia coli encodes a sensor-regulator protein for anaerob

A:Reference number: J00295; MUID:90355832; PMID:2201868

A:Accession: J00295

A:Molecule type: DNA

A:Residues: 1-468; 769-776 <IUC>

A:Cross-references: EMBL:X5315; NID:g40950; PIDN:CAA37397.1; PID:g40951

C:Genetics:

A:Gene: arcB

A:Map position: 69.5 min

C:Superfamily: aerobic respiration control sensor protein arcB; response regulator homol

C:Keywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase; s

F:23-50/Domain: transmembrane #status predicted <TM1>

F:59-77/Domain: transmembrane #status predicted <TM2>

F:78-776/Domain: intracellular #status predicted <INT>

F:526-637/Domain: response regulator homology <RR>

F:292/Binding site: phosphate (His) (covalent) #status predicted

F:574/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match

Best Local Similarity 41.5%; Score 54; DB 1; Length 776;

Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 3 WVRVRVRVRVRVRVRVR 24

DB 751 WIEEMKEWRHDEVLKAVAK 772

RESULT 5

A91140 aerobic respiration sensor-response protein [imported] - Escherichia coli (strain O157:H7

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C/Accession: A91140

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G

gawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A:Reference number: A93629; MUID:21156231; PMID:11258796

A:Accession: A91140

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-778 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA837512.1; PID:g13363562; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: EcgA089

C:Superfamily: aerobic respiration control sensor protein arcB; response regulator homo

Query Match

Best Local Similarity 41.5%; Score 54; DB 2; Length 778;

Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 3 WVRVRVRVRVRVRVRVR 24

DB 753 WIEEMKEWRHDEVLKAVAK 774

RESULT 6

D85985

aerobic respiration sensor-response protein [imported] - Escherichia coli (strain O157:H7

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001

C/Accession: D85985

R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew-

iller, L.; Grobbeck, E.J.; Davis, N.W.; Linn, A.; Dimmlanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D85985

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-778 <STO>

A:Cross-references: GB:AE005174; NID:g12517831; PIDN:AAQ58344.1; GSPDB:GN00145; UMGD:245

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: arcB

C:Superfamily: aerobic respiration control sensor protein arcB; response regulator homol

Query Match

Best Local Similarity 41.5%; Score 54; DB 2; Length 778;

Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 3 WVRVRVRVRVRVRVRVR 24

DB 753 WIEEMKEWRHDEVLKAVAK 774

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:48:06 ; Search time 9.95745 Seconds
(without alignments)
231.709 Million cell updates/sec

Title: US-10-079-075-10

Sequence: 1 RRVRRRRRRRRRRRRRRRRRRR 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	47.7	396	2 I58168	growth factor arg3
2	57.5	44.2	111	2 T14306	glycine-rich prote
3	55	42.3	62	2 F90177	hypothetical prote
4	54	41.5	776	1 R9ECAR	aerobic respiratio
5	54	41.5	778	2 A91140	aerobic respiratio
6	54	41.5	778	2 D85985	aerobic respiratio
7	53.5	41.2	105	2 H72708	probable ribosomal
8	53.5	41.2	1538	2 H70846	hypothetical glyci
9	51	39.2	432	2 A34413	atrial gland granu
10	51	39.2	570	2 T32743	hypothetical prote
11	51	39.2	778	2 AD0432	aerobic respiratio
12	51	39.2	957	2 D70835	hypothetical glyci
13	50	38.5	1145	2 A59251	myosin - Acetabula
14	49	37.7	493	2 T06031	hexokinase homolo
15	49	37.7	640	2 T08179	LRG5 protein - chl
16	49	37.7	778	2 AG0506	aerobic respiratio
17	48.5	37.3	213	2 S73607	uridine kinase udk
18	48.5	37.3	214	2 AB2157	hypothetical prote
19	48.5	37.3	516	1 F71341	probable ribose/ga
20	48.5	37.3	1210	2 A83306	hypothetical prote
21	48	36.9	114	2 T17699	arginine-rich prot
22	48	36.9	584	2 T49320	related to heterok
23	48	36.9	3511	2 A59295	unconventional myo
24	47.5	36.5	299	2 A26512	carb protein - Str
25	47.5	36.5	344	2 T34835	probable transfera
26	47.5	36.5	1007	2 H72734	hypothetical prote
27	47.5	36.5	1146	2 B70723	probable mmp12 pr
28	47	36.2	262	2 C91111	hypothetical prote
29	47	36.2	281	2 F85956	probable transpos

30	47	36.2	432	2 T05236	hypothetical prote
31	46.5	35.8	141	2 H85217	hypothetical prote
32	46.5	35.8	179	2 T05810	hypothetical prote
33	46.5	35.8	845	2 T34064	hypothetical prote
34	46	35.4	72	1 RSBPXL	excisionase - phag
35	46	35.4	72	2 S06533	excisionase - phag
36	46	35.4	72	2 A90729	excisionase [impor
37	46	35.4	93	2 AH3590	hypothetical prote
38	46	35.4	123	2 H72698	hypothetical prote
39	46	35.4	329	2 T49637	hypothetical prote
40	46	35.4	470	1 S56565	hypothetical 53K p
41	46	35.4	470	2 P91291	probable regulator
42	46	35.4	470	2 H86132	probable regulator
43	46	35.4	604	2 S60182	ATP-binding transp
44	46	35.4	604	2 H95974	msbA-like sacchari
45	46	35.4	1650	2 S28721	hypothetical prote

ALIGNMENTS

```

RESULT 1
158168
Growth factor arg3.1 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #ext_change 04-Mar-2000
C/Accession: I58168; 159386
R/Lyford, G.L.; Yamagata, K.; Kaufmann, W.E.; Barnes, C.A.; Sanders, L.K.; Copeland, N.
Neuron 14, 433-445, 1995
A/Title: Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-
A/Reference number: I58168; MUID:95161073; PMID:7857651
A/Accession: I58168
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-396 <RES>
A/Cross-references: EMBL:U09866; NID:g644828; PIDN:AAA6695.1; PID:g644829
R/Link, W.; Konietzko, U.; Kauselmann, G.; Krug, M.; Schwabe, B.; Frey, U.; Kuhl, D.
Proc. Natl. Acad. Sci. U.S.A. 92, 5734-5738, 1995
A/Title: Somatodendritic expression of an immediate early gene is regulated by synaptic
A/Reference number: I59386; MUID:95296386; PMID:777577
A/Accession: I59386
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-208, 'V', 210-396 <RE2>
A/Cross-references: EMBL:Z46925; NID:g854413; PIDN:CAA87033.1; PID:g854414
C/Genetics:
A/Supfamily: rat growth factor arg3.1

Query Match
Best Local Similarity 47.8%; Score 62; DB 2; Length 396;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 2 RRVRRRRRRRRRRRRRRRRRRR 24
DB 107 RRVKRMHWRRVFFYLRWADR 129

RESULT 2
T14306
glycine-rich protein - carrot (fragment)
C/Species: Daucus carota (carrot)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #ext_change 20-Sep-1999
C/Accession: T14306
R/Lin, X.; Huang, G.J.; Zimmerman, J.L.
submitted to the EMBL Data Library, January 1996
A/Description: Isolation and characterization of a diverse set of genes from carrot soma
A/Reference number: Z17968
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-111 <LIN>
A/Cross-references: EMBL:U47097; NID:g1276970; PID:g1276971

```


EARLIER FILING DATE: 1998-07-21
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 7
LENGTH: 1244
TYPE: PRT
ORGANISM: Rattus rattus
US-09-356-952-7

Query Match 59.0%; Score 36; DB 3; Length 1244;
Best Local Similarity 50.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVVRVVRVRR 12
DB 924 RVINVLRRHWYK 935

RESULT 14
US-08-093-453B-2
Sequence 2, Application US/08093453B
Patent No. 5439814
GENERAL INFORMATION:
APPLICANT: Frey, Teryl K.
APPLICANT: Dominguez, Geraldina
APPLICANT: Wang, Chin Yen
TITLE OF INVENTION: Modified Infectious Rubella Virus
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jamie L. Greene, Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: United States
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50
COMPUTER: Macintosh
OPERATING SYSTEM: 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,453B
FILING DATE: 19 JUL 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 07/722,334
FILING DATE: 28 JUN 1991
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 07362-0101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404 818-3700
TELEFAX: 404 818-3799
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2205 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Rubella virus
STRAIN: Thertien
US-08-093-453B-2

Query Match 59.0%; Score 36; DB 1; Length 2205;
Best Local Similarity 77.8%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRVRRRV 10
DB 103 VARVRRRV 111

RESULT 15
US-09-320-878-6
Sequence 6, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
CURRENT FILING DATE: 1999-05-27
EARLIER APPLICATION NUMBER: CIP OF 09/141,908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846,247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/119,139
EARLIER FILING DATE: 1999-02-08
EARLIER APPLICATION NUMBER: 60/100,880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087,080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 6
LENGTH: 379
TYPE: PRT
ORGANISM: Streptomyces venezuelae
FEATURE:
NAME/KEY: 251
LOCATION: unsure
OTHER INFORMATION: unsure of amino acid at this position
US-09-320-878-6

Query Match 57.4%; Score 35; DB 3; Length 379;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RVVRVVRVRR 12
DB 364 RVIDAVREWAER 375

Search completed: June 9, 2003, 12:05:08
Job time: 5.08511 secs

TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL
TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S
FILE REFERENCE: 03715.0032
CURRENT APPLICATION NUMBER: US/09/126,420A
CURRENT FILING DATE: 1998-07-30
PRIOR APPLICATION NUMBER: 60/054,351
PRIOR FILING DATE: 1997-07-31.
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 524
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-126-420A-24

Query Match 59.0%; Score 36; DB 4; Length 524;
Best Local Similarity 77.8%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VVRVRRV 10
Db 29 VVRVRRV 37

RESULT 11
US-08-318-831-3
Sequence 3, Application US/08318831
Patent No. 5656595
GENERAL INFORMATION:
APPLICANT: Schweighofer, Fabien
TITLE OF INVENTION: TOCQUE, Bruno
TITLE OF INVENTION: PEPTIDES HAVING A GDP EXCHANGE FACTOR
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SAID PEPTIDES,
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (EPO PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,831
FILING DATE: 19 October 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR92/04827
FILING DATE: 21-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92033-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-318-831-3

Query Match 59.0%; Score 36; DB 1; Length 666;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 RVVRVRRV 12

Db 346 RVLVLRHWVSK 357

RESULT 12
US-08-318-831-2
Sequence 2, Application US/08318831
Patent No. 5656595
GENERAL INFORMATION:
APPLICANT: Schweighofer, Fabien
TITLE OF INVENTION: TOCQUE, Bruno
TITLE OF INVENTION: PEPTIDES HAVING A GDP EXCHANGE FACTOR
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SAID PEPTIDES,
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (EPO PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,831
FILING DATE: 19 October 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR92/04827
FILING DATE: 21-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92033-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 814 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-318-831-2

Query Match 59.0%; Score 36; DB 1; Length 814;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RVVRVRRV 12
Db 494 RVLVLRHWVSK 505

RESULT 13
US-09-356-952-7
Sequence 7, Application US/09356952
Patent No. 6117663
GENERAL INFORMATION:
APPLICANT: Bortlack-Stodln, Ann
APPLICANT: Margalit, S. M.
APPLICANT: Bor-Sogli, Dafna
APPLICANT: Kuriyan, John
TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 600-1-228N
CURRENT APPLICATION NUMBER: US/09/356,952
CURRENT FILING DATE: 1999-07-19
EARLIER APPLICATION NUMBER: 60/093,631

```

: CURRENT FILING DATE: 1998-06-05
: EARLIER APPLICATION NUMBER: US 60/048,857
: EARLIER FILING DATE: 1997-06-06
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 5
: LENGTH: 476
: TYPE: prt
: ORGANISM: Helicobacter pylori
US-03-092-315-5

```

Query Match	63.9%;	Score 39;	DB 4;	Length 476
Best Local Similarity	45.5%;	Pred. No. 94;		
Matches 5;	Conservative	5;	Mismatches	1; Indels

```

RESULT 7
US-09-092-315-7
Sequence 7, Application US/09092315
Patent No. 6399317
GENERAL INFORMATION:
APPLICANT: TAYLOR, Diane E.
APPLICANT: Ge, Zhongxing
TITLE OF INVENTION: ALPHA-1, 3-FUCOSYLTRANSFERASE
FILE REFERENCE: 07254/049001
CURRENT APPLICATION NUMBER: US/09/092,315
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: US 60/048,857
EARLIER FILING DATE: 1997-06-06
NUMBER OF SEQ. ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO 7
LENGTH: 478
TYPE: PR1
ORGANISM: Helicobacter pylori
US-09-092-315-7

```

	Qy	Score 39;	DB 4;	Length 478;
Query Match	63.9%			
Best Local Similarity	45.5%			
Pred. No. 95;				
Matches 5; Conservative	5;	Mismatches	1;	Indels 0;
Gaps	0;			
	2 VVRVRRKWR 12			
	:::: :: :::			
Db	465 LIRAIRRWVKK 475			

```

RESULT 8
US-09-134-001C-5395
; Sequence 5395, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5395
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-5395

```

Query Match 60.7%; Score 37; DB 4; Length 193;

QY 2 VVVVVRRRV 10
| : | : | | | |
Db 157 VVVVVRRRV 165

Best Local Similarity 66.7%; Pred. No. 79;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

RESULT 9
 US-08-318-831-4
 Sequence 4, Application US/08318831
 Patent No. 5556595
 GENERAL INFORMATION:
 APPLICANT: Schweighoffer, Fabien
 APPLICANT: Tocque, Bruno
 TITLE OF INVENTION: PEPTIDES HAVING A GDP EXCHANGE FACTOR
 TITLE OF INVENTION: ACTIVITY NUCLEIC ACID SEQUENCES CODING FOR SAID PEPTIDES
 TITLE OF INVENTION: PREPARATION AND UTILIZATION
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.1 (EPO Patentin)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/318, 831
 FILING DATE: 19 October 1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR92/04827
 FILING DATE: 21-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, Julie K.
 REGISTRATION NUMBER: P-38, 619
 REFERENCE/DOCKET NUMBER: ST92033-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3839
 TELEFAX: (610) 454-3808
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 489 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 IS-08-318-831-4

Query Match 59.0%; Score 36; DB 1; Length 489;
Best Local Similarity 50.0%;
Pred. No. 2.7e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

RESULT 10
 US-09-126-420A-24
 ; Sequence 24, Application US/09126420A
 ; Patent No. 6376753
 ; GENERAL INFORMATION:
 ; APPLICANT: BATARD, YANNICK
 ; APPLICANT: ROBINEAU, TIBORCE
 ; APPLICANT: DURST, FRANCIS
 ; APPLICANT: MERCK-REICHART, DANIELE
 ; APPLICANT: DIDIERJEAN, LUC
 ; TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS
 ; TITLE OF INVENTION: TUBEROUS AND ITS APPLICATIONS AS BIOCATALYST IN

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/077,797A
FILING DATE: 14-JUN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/012,566
FILING DATE: 02-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRL276P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
FAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-077-797A-48

Query Match 63.9%; Score 39; DB 1; Length 16;
Best Local Similarity 60.0%; Pred. No. 3.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRWVR 11
DB 2 VIKMIRWVR 11

RESULT 3
PCT-US94-01238-48
Sequence 48, Application PC/TUS9401238
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
TITLE OF INVENTION: BINDING SITES AND COMPOSITIONS THEREOF
NUMBER OF SEQUENCES: 65
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01238
FILING DATE: 01-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,797
FILING DATE: 14-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/012,566
FILING DATE: 02-FEB-1993
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
PCT-US94-01238-48

Query Match 63.9%; Score 39; DB 5; Length 16;
Best Local Similarity 60.0%; Pred. No. 3.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRWVR 11
DB 2 VIKMIRWVR 11

DB 2 VIKMIRWVR 11
RESULT 4
US-09-092-315-3
Sequence 3, Application US/09092315
Patent No. 6399337
GENERAL INFORMATION:
APPLICANT: Taylor, Diane E.
TITLE OF INVENTION: ALPHA-1, 3-FUCOSYLTRANSFERASE
FILE REFERENCE: 07254/049001
CURRENT APPLICATION NUMBER: US/09/092,315
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: US 60/048,857
EARLIER FILING DATE: 1997-06-06
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 440
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-092-315-3

Query Match 63.9%; Score 39; DB 4; Length 440;
Best Local Similarity 45.5%; Pred. No. 87;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRWVR 12
DB 427 LIRAIRRWYK 437

RESULT 5
US-09-092-315-1
Sequence 1, Application US/09092315
Patent No. 6399337
GENERAL INFORMATION:
APPLICANT: Taylor, Diane E.
TITLE OF INVENTION: ALPHA-1, 3-FUCOSYLTRANSFERASE
FILE REFERENCE: 07254/049001
CURRENT APPLICATION NUMBER: US/09/092,315
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: US 60/048,857
EARLIER FILING DATE: 1997-06-06
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 464
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-092-315-1

Query Match 63.9%; Score 39; DB 4; Length 464;
Best Local Similarity 45.5%; Pred. No. 92;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRWVR 12
DB 451 LIRAIRRWYK 461

RESULT 6
US-09-092-315-5
Sequence 5, Application US/09092315
Patent No. 6399337
GENERAL INFORMATION:
APPLICANT: Taylor, Diane E.
TITLE OF INVENTION: ALPHA-1, 3-FUCOSYLTRANSFERASE
FILE REFERENCE: 07254/049001
CURRENT APPLICATION NUMBER: US/09/092,315

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:50:31 ; Search time 4.08511 Seconds
(without alignments)
86.430 Million cell updates/sec

Title: US-10-079-075-9

Perfect score: 61

Sequence: 1 RVRRVRRVRR 12

Scoring table: BLAST62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	70.5	454	4	US-09-092-315-8
2	39	63.9	16	1	US-08-077-797A-48
3	39	63.9	16	5	PCT-US94-01238-48
4	39	63.9	440	4	US-09-092-315-3
5	39	63.9	464	4	US-09-092-315-1
6	39	63.9	476	4	US-09-092-315-5
7	37	60.7	193	4	US-09-092-315-7
8	37	60.7	193	4	US-09-134-001C-5395
9	36	59.0	489	1	US-08-318-831-4
10	36	59.0	524	1	US-09-126-420A-24
11	36	59.0	666	1	US-08-318-831-3
12	36	59.0	814	1	US-08-318-831-2
13	36	59.0	1244	3	US-09-356-952-7
14	36	59.0	2205	1	US-08-093-453B-2
15	35	57.4	379	3	US-09-320-478-6
16	35	57.4	430	2	US-09-105-537-16
17	35	57.4	430	2	US-08-290-731C-9
18	35	57.4	501	4	US-09-172-952-15
19	35	57.4	841	4	US-09-413-814-107
20	35	57.4	1572	2	US-08-230-731C-5
21	35	57.4	1596	3	US-09-356-952-3
22	35	57.4	3782	4	US-09-105-537-4
23	34	55.7	38	1	US-08-179-632-7
24	34	55.7	38	1	US-08-440-174A-7
25	34	55.7	38	5	PCT-US95-00062-7
26	34	55.7	66	1	US-07-938-188-2
27	34	55.7	66	1	US-08-306-062-2

28	34	55.7	334	1	US-08-287-442-9	Sequence 9, Appli
29	34	55.7	334	1	US-08-459-701-9	Sequence 9, Appli
30	34	55.7	334	1	US-08-460-298-9	Sequence 9, Appli
31	34	55.7	334	1	US-08-459-174-9	Sequence 9, Appli
32	34	55.7	498	4	US-09-172-952-26	Sequence 26, Appli
33	33	54.1	78	4	US-08-905-223-416	Sequence 416, App
34	33	54.1	423	2	US-08-290-731C-10	Sequence 10, Appli
35	33	54.1	423	2	US-08-290-731C-11	Sequence 11, Appli
36	33	54.1	1025	2	US-08-304-309-2	Sequence 2, Appli
37	33	54.1	1025	2	US-08-304-309-4	Sequence 2, Appli
38	33	54.1	1025	3	US-08-991-942-2	Sequence 2, Appli
39	33	54.1	1025	3	US-09-138-103-2	Sequence 2, Appli
40	33	54.1	1025	4	US-09-138-103-2	Sequence 2, Appli
41	33	54.1	1025	5	PCT-US95-04567-2	Sequence 2, Appli
42	33	54.1	1025	5	PCT-US95-04567-4	Sequence 2, Appli
43	33	54.1	1297	2	US-08-290-731C-4	Sequence 2, Appli
44	33	54.1	1319	2	US-08-290-731C-2	Sequence 2, Appli
45	33	54.1	1333	3	US-09-356-952-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-092-315-8
; Sequence 8, Application US/09092315
; Patent No. 6399337
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane E.
; TITLE OF INVENTION: ALPHA-1, 3-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254/049001
; CURRENT APPLICATION NUMBER: US/09/092,315
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: US 60/048,857
; EARLIER FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-092-315-8

Query Match          70.5%; Score 43; DB 4; Length 454;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      2 VRRVRRVRR 12
Db      444 LRRVRRVWK 454

RESULT 2
US-08-077-797A-48
; Sequence 48, Application US/08077797A
; Patent No. 5679548
; GENERAL INFORMATION:
; APPLICANT: Barbae, Carlos F.
; APPLICANT: Rosenblum, Jonathan
; APPLICANT: Lerner, Richard A.
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
; TITLE OF INVENTION: BINDING SITES AND COMPOSITIONS THEREOF
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5679548th Torrey Pines Road, TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
```

```

: Sequence 11, Application PC/TUS0204812
: GENERAL INFORMATION:
:
: APPLICANT: Ronald C. Montelaro
: APPLICANT: Timothy A. Miesznar
: TITLE OF INVENTION: VIROS DERIVED ANTIMICROBIAL PEPTIDES
: FILE REFERENCE: A34001-PCT / 072396_0223
: CURRENT APPLICATION NUMBER: PCT/US02/04812
: CURRENT FILING DATE: 2002-02-19
: NUMBER OF SEQ. ID NOS: 12
: SOFTWARE: PasteSeq for Windows Version 3.0
: SEQ ID NO 11
: LENGTH: 36
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-11

```

Query Match	100.0%	Score 61	DB 1	Length 36
Best Local Similarity	100.0%	Pred. No. 0.033		
Matches 12; Conservative	0	Mismatches	0	Indels 0; Gaps 0;

```
Qy      1 RVVRVVRVRR 12
         |||||
Db      7 RVVRVVRVRR 18
```

RESULT 13
US-09-785

; Sequence 11, Application US/09785058
; GENERAL INFORMATION:

```

/ APPLICANT: Ronald C. Montelazo
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A 34001 / 072396.0222
/ CURRENT APPLICATION NUMBER: US/09/785,058
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ. ID NOS.: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
/ OS-09-785-058-11

```

Query Match	100.0%;	Score 61;	DB 21;	Length 36;
Best Local Similarity	100.0%;	Pred. No. 0.033;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	1	R V R V R V R W R R	12
Db	7	R V R V R R W R R	18

RESULT 14
US-09-785

```

; Sequence 11, Application US/09785059
; GENERAL INFORMATION:

```

1 APPLICANT: Ronald C. Monelaro
 2 APPLICANT: Timothy A. Metzner
 3 TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
 4 FILE REFERENCE: A33577 / 072396, 0217
 5 CURRENT APPLICATION NUMBER: US/09/785, 059
 6 CURRENT FILING DATE: 2001-02-16

```

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0.0

```

```

; SEQ ID NO 1
; LENGTH: 3

```

```

; TYPE: PRT
; ORGANISM: Artificial sequence

```

FEATURE:

OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-11

Query Match	100.0%	Score 61;	DB 21;	length 36;
Best Local Similarity	100.0%	Pred. No. 0.033;		
Matches	12;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	1	RVRVRVRVRVR	12
Db	7	RVRVRVRVR	18

RESULT 15

US-10-079-075-11
; Sequence 11, Application US/10079075

APPLICANT: Ronald C. Montelaro
 APPLICANT: Timothy A. Metzner
 TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
 FILE REFERENCE: A34001-A / 072396.0222
 CURRENT APPLICATION NUMBER: US/10/079,075
 CURRENT FILING DATE: 2002-02-19
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 11
 LENGTH: 36
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Artificial peptide derived from HIV-1
 JS-10-079-075-11

Query Match	100.0%;	Score 61;	DB 24;	Length 36;
Best Local Similarity	100.0%;	Pred. No. 0.033;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	12
Qy	R V R V R R V R R	
Db	7	18

Search completed: June 9, 2003, 12:25:29
Job time : 52.9574 secs

```
PCT-US02-04812-10
; Sequence 10, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-10

Query Match          100.0%; Score 61; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVRVVRVRWRR 12
    |||||
Db 13 RVRVVRVRWRR 24

RESULT 8
US-09-785-058-10
; Sequence 10, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-10

Query Match          100.0%; Score 61; DB 21; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVRVVRVRWRR 12
    |||||
Db 13 RVRVVRVRWRR 24

RESULT 9
US-09-785-059-10
; Sequence 10, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-10

Query Match          100.0%; Score 61; DB 21; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVRVVRVRWRR 12
    |||||
Db 13 RVRVVRVRWRR 24

RESULT 10
US-10-079-075-10
; Sequence 10, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-10

Query Match          100.0%; Score 61; DB 24; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVRVVRVRWRR 12
    |||||
Db 13 RVRVVRVRWRR 24

RESULT 11
PCT-US02-04432-11
; Sequence 11, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04432-11

Query Match          100.0%; Score 61; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVRVVRVRWRR 12
    |||||
Db 7 RVRVVRVRWRR 18

RESULT 12
PCT-US02-04812-11
```



```
RESULT 2
PCT-US02-04812-9
; Sequence 9, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-9

Query Match          100.0%; Score 61; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVVRVVRWVR 12
Db 1 RVVRVVRWVR 12

RESULT 3
US-09-785-058-9
; Sequence 9, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-9

Query Match          100.0%; Score 61; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVVRVVRWVR 12
Db 1 RVVRVVRWVR 12

RESULT 4
US-09-785-059-9
; Sequence 9, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: PRT
```

```
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-9

Query Match          100.0%; Score 61; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVVRVVRWVR 12
Db 1 RVVRVVRWVR 12

RESULT 5
US-10-079-075-9
; Sequence 9, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-9

Query Match          100.0%; Score 61; DB 24; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVVRVVRWVR 12
Db 1 RVVRVVRWVR 12

RESULT 6
PCT-US02-04432-10
; Sequence 10, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04432-10

Query Match          100.0%; Score 61; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVVRVVRWVR 12
Db 13 RVVRVVRWVR 24

RESULT 7
```



```
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-058-12

Query Match          100.0%; Score 61; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RVRVVRVRVR 12
        |||||
        1 RVRVVRVRVR 12

Db

RESULT 13
US-09-785-059-4
; Sequence 4, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A3577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-4

Query Match          77.0%; Score 47; DB 9; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RVRVVRVRVR 12
        |||||
        1 RVRVVRVRVR 12

Db

RESULT 14
US-10-079-075-4
; Sequence 4, Application US/10079075
; Publication No. US2002018102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-4

Query Match          77.0%; Score 47; DB 9; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RVRVVRVRVR 12
        |||||
        1 RVRVVRVRVR 12

Db

RESULT 15
US-09-785-058-4
; Sequence 4, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-4

Query Match          77.0%; Score 47; DB 9; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RVRVVRVRVR 12
        |||||
        1 RVRVVRVRVR 12

Db

Search completed: June 9, 2003, 12:34:10
Job time : 6.38298 secs
```

```
FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-11
```

```
Query Match      100.0%; Score 61; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRVRRVRRVRR 12
         |||||
Db       7 RRVRRVRRVRR 18
```

RESULT 8

```
US-10-079-075-11
Sequence 11, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-11
```

```
Query Match      100.0%; Score 61; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRVRRVRRVRR 12
         |||||
Db       7 RRVRRVRRVRR 18
```

RESULT 9

```
US-09-785-058-11
Sequence 11, Application US/09785058
Publication No. US20030036627A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A 34001 / 072396.0222
CURRENT APPLICATION NUMBER: US/09/785.058
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-11
```

```
Query Match      100.0%; Score 61; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRVRRVRRVRR 12
         |||||
Db       7 RRVRRVRRVRR 18
```

RESULT 10

```
US-09-785-059-12
Sequence 12, Application US/09785059
Patent No. US20020169279A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-059-12
```

```
Query Match      100.0%; Score 61; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRVRRVRRVRR 12
         |||||
Db       1 RRVRRVRRVRR 12
```

RESULT 11

```
US-10-079-075-12
Sequence 12, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: artificial peptides derived from HIV-1
US-10-079-075-12
```

```
Query Match      100.0%; Score 61; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRVRRVRRVRR 12
         |||||
Db       1 RRVRRVRRVRR 12
```

RESULT 12

```
US-09-785-058-12
Sequence 12, Application US/09785058
Publication No. US20030036627A1
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-9
```

```
Query Match          100.0%; Score 61; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RVVRVVRWVR 12
Db      1 RVVRVVRWVR 12
```

```
RESULT 3
US-09-785-058-9
; Sequence 9, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
```

```
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-9
```

```
Query Match          100.0%; Score 61; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RVVRVVRWVR 12
Db      1 RVVRVVRWVR 12
```

```
RESULT 4
US-09-785-059-10
; Sequence 10, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
```

```
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-10
```

```
Query Match          100.0%; Score 61; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RVVRVVRWVR 12
Db      13 RVVRVVRWVR 24
```

```
RESULT 5
US-10-079-075-10
; Sequence 10, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
```

```
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-10
```

```
Query Match          100.0%; Score 61; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RVVRVVRWVR 12
Db      13 RVVRVVRWVR 24
```

```
RESULT 6
US-09-785-058-10
; Sequence 10, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
```

```
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-10
```

```
Query Match          100.0%; Score 61; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RVVRVVRWVR 12
Db      13 RVVRVVRWVR 24
```

```
RESULT 7
US-09-785-059-11
; Sequence 11, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
```

```
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 12:01:36 ; Search time 6.38298 Seconds
(without alignments)
194.092 Million cell updates/sec

Title: US-10-079-075-9
Perfect score: 61
Sequence: 1 RVKRVRRWRR 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	12	9	US-09-785-059-9
2	61	100.0	12	9	US-10-079-075-9
3	61	100.0	12	9	US-09-785-058-9
4	61	100.0	24	9	US-09-785-059-10
5	61	100.0	24	9	US-10-079-075-10
6	61	100.0	24	9	US-09-785-058-10
7	61	100.0	36	9	US-09-785-059-11
8	61	100.0	36	9	US-10-079-075-11
9	61	100.0	36	9	US-09-785-058-11
10	61	100.0	48	9	US-09-785-059-12
11	61	100.0	48	9	US-10-079-075-12
12	61	100.0	48	9	US-09-785-058-12
13	61	100.0	12	9	US-09-785-059-4
14	61	100.0	12	9	US-10-079-075-4
15	61	100.0	12	9	US-09-785-058-4
16	61	100.0	24	9	US-09-785-059-5
17	61	100.0	24	9	US-10-079-075-5
18	61	100.0	24	9	US-09-785-058-5
19	61	100.0	24	9	US-09-785-059-6

20	47	77.0	36	9	US-10-079-075-6	Sequence 6, Appl1
21	47	77.0	36	9	US-09-785-058-6	Sequence 6, Appl1
22	47	77.0	42	9	US-09-785-059-7	Sequence 7, Appl1
23	47	77.0	42	9	US-10-079-075-7	Sequence 7, Appl1
24	47	77.0	42	9	US-09-785-058-7	Sequence 7, Appl1
25	47	77.0	48	9	US-09-785-059-8	Sequence 8, Appl1
26	47	77.0	48	9	US-10-079-075-8	Sequence 8, Appl1
27	47	77.0	48	9	US-09-785-058-8	Sequence 8, Appl1
28	43	70.5	454	9	US-10-120-319-8	Sequence 18, Appl1
29	43	70.5	454	10	US-09-733-524-18	Sequence 18, Appl1
30	40	65.6	46	10	US-09-864-761-48882	Sequence 18, Appl1
31	39	63.9	31	9	US-09-785-059-3	Sequence 3, Appl1
32	39	63.9	31	9	US-10-079-075-3	Sequence 3, Appl1
33	39	63.9	31	9	US-09-785-058-3	Sequence 3, Appl1
34	39	63.9	440	9	US-10-120-319-3	Sequence 3, Appl1
35	39	63.9	440	10	US-09-733-524-3	Sequence 3, Appl1
36	39	63.9	464	9	US-10-120-319-5	Sequence 3, Appl1
37	39	63.9	476	9	US-10-120-319-5	Sequence 5, Appl1
38	39	63.9	476	10	US-09-733-524-15	Sequence 15, Appl1
39	39	63.9	478	9	US-10-120-319-7	Sequence 15, Appl1
40	39	63.9	479	10	US-09-733-524-17	Sequence 17, Appl1
41	39	63.9	501	10	US-09-733-524-1	Sequence 1, Appl1
42	38	62.3	143	9	US-10-174-590-340	Sequence 340, App
43	38	62.3	143	9	US-10-176-758-340	Sequence 340, App
44	38	62.3	143	9	US-10-175-737-340	Sequence 340, App
45	38	62.3	143	9	US-10-173-706-340	Sequence 340, App

ALIGNMENTS

```
RESULT 1
US-09-785-059-9
; Sequence 9, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-9

Query Match      100.0%; Score 61; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RVKRVRRWRR 12
Db      1 RVKRVRRWRR 12

RESULT 2
US-10-079-075-9
; Sequence 9, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
```

```

OS Unidentified.
XX
XX WO200102600-A2.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000MO-US18057.
XX
XX 06-JUL-1999; 99US-0347878.
XX
XX 06-DEC-1999; 99US-0457205.
XX
XX (GEAT ) GEN ATOMICS.
XX
XX Yuan C;
XX
XX WPI; 2001-071583/08.
XX
XX Assaying method, useful for prognosis and diagnosis of disease,
XX comprises contacting sample with a mutant analyte-binding enzyme and
XX detecting binding -
XX
XX PS Disclosure; Page 133; 187pp; English.
XX
XX CC The present invention relates to a method for assaying an analyte in a
XX CC sample comprising: contacting the sample with a mutant analyte-binding
XX CC enzyme which has binding affinity for the analyte or an immediate
XX CC analyte enzymatic conversion product but has attenuated catalytic
XX CC activity; and detecting resulting binding. The method is useful in
XX CC monitoring biological systems/processes, or prognosis/diagnosis of
XX CC disease caused by imbalances of the analytes. The present sequence is a
XX CC peptide used in the present invention.
XX
XX SQ Sequence 16 AA;
XX
XX Query Match 63.9%; Score 39; DB 22; Length 16;
XX Best Local Similarity 60.0%; Pred. No. 14;
XX Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 2 VVVRVRRVVR 11
XX |:::|||||
XX 2 VIKWIRRVVR 11
XX
XX RESULT 15
XX AAY20593
XX ID AAY20593 standard; Protein; 31 AA.
XX
XX AC AAY20593;
XX
XX DT 22-JUL-1999 (first entry)
XX
XX DE Human neurofilament-L mutant protein fragment 100.
XX
XX KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
XX KW frameshift mutation; age-related disease; neurodegenerative disorder;
XX KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
XX KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
XX KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
XX KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
XX KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
XX KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
XX KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
XX KW high mobility group protein-C; neuroendocrine specific protein A.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX PN WO9845322-A2.
XX
XX PD 15-Oct-1998.
XX
XX PF 02-APR-1998; 98WO-IB00705.
XX

```

```

PR 10-APR-1997; 97US-0043163.
XX
XX (UYUT-) RIJKSUNIV UTRECHT.
XX
XX (ROYA-) ROYAL NETHERLANDS ACADEMIES OF SCI.
XX
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX
XX WPI; 1998-609901/51.
XX
XX DR N-PSDB; AAX75758.
XX
XX PT Diagnosing disease by detecting frameshift mutations in RNA or
XX PT corresponding protein mutations - used to diagnose cancer and
XX PT neurological diseases, particularly Alzheimer's disease, and also
XX PT for treatment and prevention with specific ribozymes or wild-type
XX PT RNA
XX
XX PS Disclosure; Figure 7; 258pp; English.
XX
XX CC This invention describes a novel method for the diagnosis of a disease
XX CC caused by, or associated with, an RNA molecule that has a frameshift
XX CC mutation. The method is used to diagnose age-related diseases, especially
XX CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
XX CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX CC and many others listed) or susceptibility to these disorders. The method
XX CC allows a definitive diagnosis of Alzheimer's disease in living patients,
XX CC at an early stage. It is based on the observation that disease may be
XX CC caused by mutations in RNA rather than DNA. The invention describes the
XX CC use of neuronal system RNA molecules, specifically proteins including
XX CC beta-amyloid precursor protein (beta-APP), the microtubule associated
XX CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
XX CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
XX CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
XX CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
XX CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
XX CC protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
XX SQ Sequence 31 AA;
XX
XX Query Match 63.9%; Score 39; DB 19; Length 31;
XX Best Local Similarity 63.6%; Pred. No. 27;
XX Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 2 VVVRVRRVVR 12
XX |:::|||||
XX 12 VVRFVRRVIRQ 22
XX
XX Search completed: June 9, 2003, 11:55:35
XX Job time : 14.9362 secs

```

CC useful as herbicides.

XX Sequence 662 AA;

Query Match 65.6%; Score 40; DB 23; Length 662;
Best Local Similarity 70.0%; Pred. No. 4e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRWVR 11
|:::|||||
Db 596 LVRVRRWVR 605

RESULT 12

AA87952
ID AA87952 standard; peptide; 16 AA.

XX AA87952;

DT 28-MAR-1995 (first entry)

DE Randomly generated iron (III) chloride-binding site Fe-7.

XX Metal cation binding site: iron (III) chloride; immunoglobulin;

KM CDR3; complementarily determining region; mutagenesis;

XX metalloantibody; phagemid display library; magnetic antibody.

OS Synthetic.

XX WO9418220-A.

XX 18-AUG-1994.

XX 02-FEB-1994; 94WO-US012238.

PR 02-FEB-1993; 93US-0012566.

XX 14-JUN-1993; 93US-0077797.

PA (SCRI) SCRIpps RES INST.

PI Barbas CF, Lerner RA, Rosenblum J;

XX WPI; 1994-279674/34.

XX Methods using oligo-nucleotide primers in prodn of metal binding

PT sites in CDR regions of immunoglobulin heavy or light-chains -

XX for use therapeutically, diagnostically or as metal ion chelators

PS Claim 26; Page 27; 141pp; English.

XX A human monoclonal antibody capable of immunoreacting with

CC iron (III) chloride has one of the heavy chain variable region

CC amino acid sequences AA87946-R57953. The binding sequences were

CC generated by mutagenic PCR on the Ig heavy chain gene and were

CC identified by expression on phagemid display proteins.

XX Sequence 16 AA;

Query Match 63.9%; Score 39; DB 15; Length 16;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRWVR 11
|:::|||||
Db 2 VIKIRRWVR 11

RESULT 13

AA878104
ID AA878104 standard; Peptide; 16 AA.

XX AA878104;
XX

DT 20-NOV-2001 (first entry)

XX Fe(III) metal ion binding sequence SEQ ID NO 73.

DE abnormal base-pairing; polymorphism; nucleic acid repair enzyme;

KM mutation; infection; cancer; immune system disorder; metabolic disorder;

KM muscle disorder; bone disorder; nervous system disorder;

XX high throughput screening.

OS Synthetic.

XX WO200162968-A2.

XX 30-AUG-2001.

XX 05-JAN-2001; 2001WO-US00452.

XX 25-FEB-2000; 2000US-0514016.

PA (GEAT) GEN ATOMICS.

PI Yuan C;

XX WPI; 2001-536643/59.

XX Detecting abnormal base-pairing, mutation in nucleic acid, or

PT polymorphism in gene locus, comprises contacting nucleic acid with

XX abnormal base-pairing and mutant nucleic acid repair enzyme, and

XX detecting their binding

XX Disclosure, Page 217; 294pp; English.

XX The invention relates to detecting abnormal base-pairing in a nucleic

CC acid duplex, mutation in a nucleic acid or polymorphism in a gene locus,

CC comprising contacting a nucleic acid duplex having an abnormal

CC base-pairing with a mutant nucleic acid repair enzyme or its complex and

CC detecting the binding between the nucleic acid duplex and mutant enzyme,

CC such that the presence of abnormal base-pairing, mutation or polymorphism

CC is detected. The method is useful for prognosis or diagnosis of the

CC presence or severity of a disease, disorder or infection by a

CC pathological agent associated with the mutation, including cancer, immune

CC system disorders, metabolic disorders, muscle and bone disorders, nervous

CC system disorders, signal disorders and transporter disease or disorder.

CC The method is rapid and accurate and is amenable to high throughput

CC formats. The method requires neither specific probes nor gel

CC electrophoresis and is amenable to automation for simultaneous

CC detection of a large number of nucleic acid mutations. The present

CC sequence is that of a metal ion binding peptide sequence, useful to the

XX invention.

XX Sequence 16 AA;

Query Match 63.9%; Score 39; DB 22; Length 16;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRWVR 11
|:::|||||
Db 2 VIKIRRWVR 11

RESULT 14

AA866859
ID AA866859 standard; Peptide; 16 AA.

XX AA866859;

DT 27-APR-2001 (first entry)

XX Metal ion binding sequence #48.

XX Analyte-binding enzyme; analyte analysis.
XX

CC and anti-infective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX
SQ Sequence 64 AA;

Query Match 65.6%; Score 40; DB 23; Length 64;
Best Local Similarity 70.0%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 3 VRVRRVRR 12
:|||||
Db 47 LRVLRMPRR 56

RESULT 10

AAU67788 standard; Protein; 103 AA.

AAU67788;

27-FEB-2002 (first entry)

Propionibacterium acnes immunogenic protein #28684.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopetrotic; neuroprotectant.

OS Propionibacterium acnes.

XX
XX WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIYA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59674.

PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID NO 28983; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.

XX
SQ Sequence 103 AA;

Query Match 65.6%; Score 40; DB 22; Length 103;
Best Local Similarity 70.0%; Pred. No. 63;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 3 VRVRRVRR 12
:|||||
Db 87 LRVLRMPRR 96

RESULT 11

ABB92719 standard; Protein; 662 AA.

ABB92719;

31-MAY-2002 (first entry)

Herbicideally active polypeptide SEQ ID NO 1930.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

PN 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicideally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
XX Claim 5; SEQ ID NO 1930; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicideally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB160511), expressed DNA
 CC sequences (AB161840-AB16175) and the encoded proteins
 CC (AB161737-AB1617072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 458 AA;

Query Match 67.2%; Score 41; DB 22; Length 458;
 Best Local Similarity 87.5%; Pred. No. 2e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVVVRRWV 10
 |||||
 91 VVVVRRWV 98

RESULT 8
 AAM38542 standard; Protein; 46 AA.
 XX
 AC AAM38542;

DT 17-OCT-2001 (first entry)

DE Peptide #12579 encoded by probe for measuring placental gene expression.

XX
 KM Probe: microarray; human; placenta; antenatal diagnosis;
 KM genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488897/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX XX analyzing gene expression in human placenta -

XX PS Claim 27; SEQ ID NO 38811; 654pp; English.

XX XX The present invention relates to single exon nucleic acid probes (SENP;
 CC see A131315-A157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX
 SQ Sequence 46 AA;

Query Match 65.6%; Score 40; DB 22; Length 46;
 Best Local Similarity 63.6%; Pred. No. 28;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 VVVVRRWVR 12
 :|||:||||
 19 IVVKGKWR 29

RESULT 9
 ABP31941
 ID ABP31941 standard; Protein; 64 AA.
 XX
 AC ABP31941;

DT 08-JUL-2002 (first entry)

DE Human ORF914 protein, SEQ ID NO:1828.

XX Human; ORF; open reading frame; ORF; drug screening; diagnosis;
 KM disease monitoring; cytokine; cell proliferation; cell differentiation;
 KM immune modulation; haematopoiesis regulation; tissue growth;
 KM angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KM thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KM behaviour; cancer; proliferative disorder; neurological disorder;
 KM cardiovascular disease; immune system disorder; organ transplantation;
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KM hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KM vasotropic; antipariatic; antidiabetic; cyostatic; noctropic;
 KM neuroprotective; antithrombotic; anticoagulant; thrombolytic;
 KM cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KM dermatological; analgesic; virucide; antibacterial; fungicide.

XX OS Homo sapiens.

XX PN WO200190366-A2.

XX PD 29-NOV-2001.

XX PF 24-MAY-2001; 2001WO-US17076.

XX PR 24-MAY-2000; 2000US-206690P.

XX PA (CURA-) CURAGEN CORP.

XX PI Leach MD, Shinkets RA;

XX XX WPI; 2002-106200/14.

XX DR N-PSDB; AEN75967.

XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -

XX XX Claim 10; Page 721; 2508pp; English.

XX XX Sequences ABP31028-ABP35561 represent 434 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences AEN75054-
 CC AEN7587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,

DR WPI; 2000-105868/09.
DR N-PSDB; AA245404.
XX
PT Novel receptor protein for screening compounds used in treating
PT Irritable bowel syndrome, constipation and other gastric conditions
XX
PS Claim 5; Fig 5; 44pp; English.
XX
CC The present sequence represents splice variant MTL-R1B of the motilin
CC receptor. The gene encodes a G-protein coupled receptor, and is
CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
CC MTL-R1A (see AA554145) and MTL-R1B (see AA554146). MTL-R1A is a
CC functional seven transmembrane domain form, and MTL-R1B is a truncated
CC five transmembrane domain. The MTL-R1 proteins are used to identify
CC agonists and antagonists which can be used for treating gastric motility
CC disorders, functional defects, disorders secondary to neurological
CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC dysmotility, diabetes, infections, stress-related motility disorders,
CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC constipation, chronic idiopathic pseudo obstruction, acute faecal
CC impaction, postoperative ileus, gallstones, infantile colic, irritable
CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
CC diarrhoea. They can also be used in the preparation for colonoscopy,
CC endoscopy and duodenal intubation.
XX
SQ Sequence 386 AA;
XX
Query Match 67.2%; Score 41; DB 21; Length 386;
Best Local Similarity 58.3%; Pred. No. 1.7e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 RYRVVRRMVR 12
Db 295 QTVRLKMSRR 306
XX
RESULT 6
AAB62653
ID AAB62653 standard; Protein; 386 AA.
XX
AC AAB62653;
XX
DT 23-JUL-2001 (first entry)
XX
DE Short form of motilin receptor, GPR-38B isoform.
XX
KW zsig33; signal transduction; hormone; enzyme; neural development;
KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
KW glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R;
KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.
XX
XX Homo sapiens.
XX OS
XX WO200138355-A2.
XX PN
XX PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US32074.
XX PF
XX 22-NOV-1999; 99US-0166765.
XX PR
XX (ZYMO) ZYMOGENETICS INC.
XX PA
XX (ZYMO) ZYMOGENETICS INC.
XX PI
XX Shepard PO, Jaspers SR, Deisher TA, Bishop PD;
XX
XX WPI; 2001-355879/37.
XX DR N-PSDB; AAF83684.
XX
PT Forming reversible peptide receptor complex for purifying cell and
PT peptides, stimulating signal transduction and modulating hormone
PT secretion, involves contacting a receptor with zsig33 polypeptide
XX

PS Disclosure; Page 106-109; 11pp; English.
XX
XX The invention relates to a method of forming a reversible peptide-
CC receptor complex that involves providing an immobilized receptor, and
CC contacting the receptor with a zsig33 peptide (comprising residues 24-37
CC of AAB62649), where the receptor binds to the zsig33 peptide. The method
CC is useful for purifying cells, purifying a peptide, stimulating signal
CC transduction in a cell expressing a receptor. It is also useful for
CC modulating secretion of hormones, neural development and/or utilization,
CC gastric contractility, nutrient uptake, secretion of digestive and
CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
CC hormone secretion in a mammal having a disease associated with abnormal
CC levels of growth hormone, such as osteoporosis, bone repair, bone
CC remodeling, low osteoblast levels, cartilage repair and remodeling,
CC skeletal dysplasia, immune suppression, obesity, growth retardation,
CC protein catabolic responses after surgery, cachexia, protein loss,
CC dwarfism, wound healing and ovulation induction, treating a mammal having
CC a metabolic disorder requiring neurological feedback, such as satiety
CC regulation, glucose absorption and metabolism and neuropathy-associated
CC gastrointestinal disorders, and stimulating glucose-induced insulin
CC release in a mammal. The present sequence represents the short form of
CC motilin receptor, GPR-38B (one of the two isoforms of GPR38 which result
CC from alternative splicing). GPR38 has homology to the human G-protein
CC coupled receptor, GHS-R.
XX
SQ Sequence 386 AA;
XX
Query Match 67.2%; Score 41; DB 22; Length 386;
Best Local Similarity 58.3%; Pred. No. 1.7e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 RYRVVRRMVR 12
Db 295 QTVRLKMSRR 306
XX
RESULT 7
AAB64714
ID AAB64714 standard; Protein; 458 AA.
XX
AC AAB64714;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 20934.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmacutical.
XX KW
XX Drosophila melanogaster.
XX OS
XX WO200171042-A2.
XX PN
XX PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX PF
XX 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX PA
XX Venter JC, Adams M, Li PWD, Myers EW;
XX PI
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL08817.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 20934; 21pp + Sequence Listing; English.

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 48268; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations in
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG0377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 201 AA;
Query Match 72.1%; Score 44; DB 22; Length 201;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 RVRVVRVRW 9
|:::|::|
Db 36 RIRIRVRW 44
RESULT 4
ID AAG65539 standard; peptide; 18 AA.
XX AAG65539;
AC
XX
DT 30-NOV-2001 (first entry)
XX
DE Peptide sequence used in the course of the invention.
XX
KM Antimicrobial; microbial membrane disrupter; gene therapy; pathogen;
KM spermicide; imaging; magatnin; PglA.
XX
OS Synthetic.
XX
PN WO200160162-A2.
XX
PD 23-AUG-2001.
XX
PF 15-FEB-2001; 2001WO-US04822.
XX
PR 15-FEB-2000; 2000US-0182495.
XX
XX (UYOH-) UNIV OHIO.
XX
XX Blazyk JF;
XX
PI WPI; 2001-565322/63.
XX
DR Novel peptides having antimicrobial activity have positive charge to
PT selectively disrupt microbial membranes, assume beta sheet structure in
PT membrane environment and are substantially amphipathic in beta sheet
PT structure -
XX
XX Example 102; Page 84; 119pp; English.

XX The invention provides an antimicrobial compound (I) which is a peptide
CC having 8-50 amino acids, a net charge of 4, a hydrophobic moment (microh)
CC as a beta sheet which is 0.2 higher than its micro H as alpha helix, and
CC having detectable membrane disrupting activity against a microbial
CC pathogen, and substantially no membrane disrupting activity against
CC mammalian cells. (I) is useful for inhibiting microbial activity.
CC has a detectable membrane disrupting activity against a pathogen, and is
CC useful for inhibiting non-microbial pathogenic activity also. (I) is also
CC useful for killing human sperm. The peptides are also provided in the
CC form of an expression vector comprising a nucleic acid encoding the
CC peptides. The peptides are useful for inhibiting the activity of
CC bacteria, and other microbial pathogens such as algae, fungi or protozoa
CC and for inhibiting non-microbial pathogens such as worms or arthropods,
CC as and spermicides for humans as the sperm membrane is atypical of human
CC cell membranes. (I) also has diagnostic uses e.g., in localizing an
CC infection or detecting sepsis. The peptides may act as binding molecules
CC and are useful to purify a target from blood, for qualitative or
CC quantitative analysis of analytes in in vitro sample, and for in vivo
CC imaging. Also, they are useful as molecular weight markers, as nutrient
CC source, as growth medium component for culturing microorganisms, as well
CC as a food ingredient for human consumption. The peptides have a greater
CC selectivity for bacterial versus mammalian lipids as compared to the
CC alpha helical peptides. Sequences AAG65536-47 represent amino acid
CC sequences of antimicrobial peptides.
SQ Sequence 18 AA;
Query Match 68.9%; Score 42; DB 22; Length 18;
Best Local Similarity 50.0%; Pred. No. 5.7;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 RVRVVRVRW 12
|:::|::|
Db 6 RIRIRIRIR 17
RESULT 5
ID AAY54146 standard; Protein; 386 AA.
XX AAY54146;
AC
XX
DT 27-MAR-2000 (first entry)
XX
DE Amino acid sequence of the motilin receptor splice variant MTL-R1B.
XX
KM Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KM spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KM functional defect; neurological disorder; scleroderma; colonoscopy;
KM paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KM infection; stress-related motility disorder; psychogenic disorder;
KM gastroparesis; gastro-oesophageal reflux disease; constipation;
KM chronic idiopathic pseudo obstruction; acute faecal impaction;
KM postoperative ileus; gallstones; infantile colic; diarrhoea;
KM irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
KM endoscopy; duodenal intubation.
XX
OS Homo sapiens.
XX
XX WO9964436-A1.
XX
PN
XX
PD 16-DEC-1999.
XX
PF 08-JUN-1999; 99WO-US12773.
XX
XX 12-JUN-1998; 98US-0089098.
XX
PA (MERI) MERCK & CO INC.
XX
PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI Pong S, Smith RG;
XX

```

DR N-PSDB,AAS59545.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 10673; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
SO Sequence 66 AA;

Query Match 78.7%; Score 48; DB 22; Length 66;
Best Local Similarity 66.7%; Pred. No. 2.8;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Cy 1 RVRRVRRRVRR 12
   ||:|||||:|
Db 1 RRIRKVRWQR 12

RESULT 2
AAU41774
ID AAU41774 standard; Protein; 144 AA.
AC AAU41774;
XX
XX DT 13-FEB-2002 (first entry)
DE Propionibacterium acnes immunogenic protein #2670.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone/joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX OS Propionibacterium acnes.
XX
XX PN WO200181581-A2.
XX
XX PD 01-NOV-2001.
XX
XX PF 20-APR-2001; 2001WO-US12865.
XX
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-200841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX SKetky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2003-616774/71.
DR
```

DR	N-PSDB; AAS59515.
XX	
PT	Propionibacterium acnes polypeptides and nucleic acids useful for
PT	vaccinating against and diagnosing infections, especially useful for
PT	treating acne vulgaris -
XX	
PS	Example 1; SEQ ID No 2969; 1069pp; English.
XX	
CC	Sequences AAU39105-AAU58017 represent Propionibacterium acnes immunogenic
CC	polypeptides. The proteins and their associated DNA sequences are used in
CC	the treatment, prevention and diagnosis of medical conditions caused by
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC	pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC	P. acnes is also involved in infections of bone, joints and the central
CC	nervous system, however it is particularly involved in the inflammatory
CC	lesions associated with acne vulgaris. A method for detecting the
CC	presence or absence of P. acnes in a patient comprises contacting a
CC	sample with a binding agent that binds to the proteins of the invention
CC	and determining the amount of bound protein in the sample. The
CC	polypeptides may be used as antigens in the production of antibodies
CC	specific for P. acnes proteins. These antibodies can be used to
CC	deregulate expression and activity of P. acnes polypeptides and
CC	therefore treat P. acnes infections. The antibodies may also be used as
CC	diagnostic agents for determining P. acnes presence, for example, by
CC	enzyme linked immunosorbent assay (ELISA).
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
SQ	
	Sequence 144 AA;
Query Match	72.1%; Score 44; DB 22; Length 144;
Best Local Similarity	75.0%; Pred. No. 23;
Matches 9; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Oy	1 RVRVRRRWVRR 12 27 RVSEVVARWVR 38
DB	
RESULT 3	
ID	ABG17909 standard; Protein; 201 AA.
AC	ABG17909;
DT	18-FEB-2002 (first entry)
XX	
DB	Novel human diagnostic protein #17900.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
PD	11-OCT-2001.
PF	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
PA	(HYSE-) HYSEQ INC.
XX	
FI	Dzmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS82096.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:35:06 ; Search time 12.9362 Seconds
(without alignments)
123.607 Million cell updates/sec

Title: US-10-079-075-9
Perfect score: 61
Sequence: 1 RVRVRVRRVRR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	78.7	66	22	AAU49478
2	44	72.1	144	22	AAU41774
3	44	72.1	201	22	ABG17909
4	42	68.9	18	22	AA665539
5	41	67.2	386	21	AA554146
6	41	67.2	386	22	AB662653
7	41	67.2	458	22	AB664714
8	40	65.6	46	22	AA38542
9	40	65.6	64	23	ABP31941
10	40	65.6	103	22	AAU67788

11	40	65.6	662	23	AB922719	Herbicidally activ
12	39	63.9	16	15	AA57952	Randomly generated
13	39	63.9	16	22	AA678104	Fel(III) metal ion
14	39	63.9	16	22	AA66859	Metal ion binding
15	39	63.9	31	19	AA20593	Human neurofilamen
16	39	63.9	50	20	AA89340	Helicobacter pylori
17	39	63.9	104	22	AA62267	Propionibacterium
18	39	63.9	464	20	AA86008	Helicobacter pylori
19	38	62.3	9	23	AAU99184	Partial human pert
20	38	62.3	91	22	AAU39112	Propionibacterium
21	38	62.3	104	22	AA84315	Human immune/haema
22	38	62.3	142	21	AA641902	Human immune/thalia
23	38	62.3	143	21	AA99423	Human PRO1482 (UNQ
24	38	62.3	143	22	AAU29193	Human PRO polypept
25	38	62.3	143	22	AA666172	Protein of the inv
26	38	62.3	261	21	AA641901	Arabidopsis thalia
27	38	62.3	280	21	AA641900	Arabidopsis thalia
28	38	62.3	715	20	AA705525	HIV-1 group O isol
29	38	62.3	745	22	AB66409	Drosophila melanog
30	37	60.7	48	21	AA613632	Arabidopsis thalia
31	37	60.7	48	21	AA613632	Arabidopsis thalia
32	37	60.7	104	20	AA712551	Human 5' EST seque
33	37	60.7	141	21	AA609943	Arabidopsis thalia
34	37	60.7	193	23	AA640550	Staphylococcus epi
35	37	60.7	237	22	AAU27640	Human protein AFP6
36	37	60.7	284	22	AAU01781	Human secreted pro
37	37	60.7	332	20	AA736034	Extended human sec
38	37	60.7	496	22	AA681335	Human AFP protein
39	37	60.7	496	22	AB88466	Human membrane or
40	37	60.7	497	21	AA812138	Hydrophobic domain
41	37	60.7	510	21	AA643982	Human cancer assoc
42	37	60.7	1357	22	AA639416	Human polypeptide
43	37	60.7	1384	22	AA639418	Human polypeptide
44	37	60.7	1410	22	AA639417	Human polypeptide
45	37	60.7	1415	22	AAU30918	Novel human secret

ALIGNMENTS

RESULT 1						
ID	AAU49478	standard;	Protein;	66	AA.	
XX	XX					
AC	AAU49478;					
DT	27-FEB-2002	(first entry)				
DE	Propionibacterium	acnes immunogenic protein #10374.				
XX	XX					
XX	SAPHO syndrome;	synovitis; acne; pustulosis; hypertosis; osteomyelitis;				
KW	uvellitis; endophthalmitis;	bone; joint; central nervous system; ELISA;				
KW	inflammatory lesion; acne vulgaris;	enzyme linked immunosorbent assay;				
KW	dermatological; osteopneic;	neuroprotectant.				
OS	Propionibacterium	acnes.				
XX	XX					
PN	MO200181581-A2.					
XX	XX					
PD	01-NOV-2001.					
XX	XX					
PF	20-APR-2001;	2001WO-US12865.				
XX	XX					
PR	21-APR-2000;	2000US-199047P.				
PR	02-JUN-2000;	2000US-208841P.				
PR	07-JUL-2000;	2000US-216747P.				
XX	XX					
PA	(CORI-) CORIXA CORP.					
XX	XX					
PI	Skelly YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;					
PI	L-malsonneuve J, Zhang Y, Jen S, Carter D;					
XX	XX					
DR	WPI; 2001-616774/71.					

RESULT 15

ID 09MT99 PRELIMINARY; PRT; 476 AA.
 AC 09MT99;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE Probable intron maturase (Maturase K) (Fragment).
 GN MATK
 OS Cryptomeria fortunei.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Cupressaceae; Cryptomeria.
 OX NCBI_TaxID=99810;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Kusumi J., Tsunuma Y., Yoshimaru H., Tachida H.;
 RT "Phylogenetic relationships in Taxodiaceae and Cupressaceae based on
 the matk, chl, trnL-trnF IGS region and trnL intron sequences."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases
 CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
 CC INTRONS (BY SIMILARITY).
 CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS, BY
 CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
 CC MITOCHONDRIAL INTRONS.
 CC EMBL; AB030117; BAB01546.1; -
 DR InterPro; IPR000442; Intron_mature2.
 DR InterPro; IPR002866; MatK_N.
 DR Pfam; PF01348; Intron_mature2; 1.
 DR Pfam; PF01824; MatK_N; 1.
 KM Chloroplast; mRNA processing.
 FT NON TER 476
 SQ SEQUENCE 476 AA; 56792 MM; 7E7D491074C15FB2 CRC64;

Query Match 67.2%; Score 41; DB 8; Length 476;
 Best Local Similarity 60.0%; Pred. No. 77;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VVRVRRVR 11
 :||:||||
 Db 159 LVRIFFRRIR 168

Search completed: June 9, 2003, 12:01:09
 Job time : 12.0213 secs

SO SEQUENCE 470 AA; 56225 MW; 2884AD1B78A7619 CRC64;

Query Match 67.2%; Score 41; DB 8; Length 470;

Best Local Similarity 60.0%; Pred. No. 76;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VWRVRRWR 11

DB 119 LVRIFFRRWR 128

RESULT 12

Q9MT98 PRELIMINARY; PRT; 471 AA.

AC Q9MT98; 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)

GN Probable intron maturase (Maturase K) (Fragment).

OS Glycostrobus pensilis.

OC Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;

OC Glycostrobus.

OX NCBI_TaxID=99811;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LEAF;

RA Kusumi J., Tsunura Y., Yoshimaru H., Tachida H.;

RT "Phylogenetic relationships in Taxodiaceae and Cupressaceae based on the matK, chlL, trnL-trnF IGS region and trnL intron sequences."

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II

CC INTRONS (BY SIMILARITY).

CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,

CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY

CC MITOCHONDRIAL INTRONS.

CC EMBL; AB030118; BAB01547.1; -.

DR InterPro: IPR000442; Intron_maturase2.

DR InterPro: IPR002866; MatK_N

DR Pfam; PF01348; Intron_maturase2; 1.

KW Chloroplast; mRNA processing.

FT NON TER 471

SO SEQUENCE 471 AA; 56153 MW; 6DAB020074EF2EB4 CRC64;

QY Query Match 67.2%; Score 41; DB 8; Length 471;

Best Local Similarity 60.0%; Pred. No. 76;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

RA Kusumi J., Tsunura Y., Yoshimaru H., Tachida H.;

RT "Phylogenetic relationships in Taxodiaceae and Cupressaceae based on the matK, chlL, trnL-trnF IGS region and trnL intron sequences."

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II

CC INTRONS (BY SIMILARITY).

CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,

CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY

CC MITOCHONDRIAL INTRONS.

CC EMBL; AB030121; BAB01550.1; -.

DR InterPro: IPR000442; Intron_maturase2.

DR InterPro: IPR002866; MatK_N

DR Pfam; PF01348; Intron_maturase2; 1.

KW Chloroplast; mRNA processing.

FT NON TER 476

SO SEQUENCE 476 AA; 56628 MW; 80D45EA69F20DD46 CRC64;

QY Query Match 67.2%; Score 41; DB 8; Length 476;

Best Local Similarity 60.0%; Pred. No. 77;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VWRVRRWR 11

DB 159 LVRIFFRRWR 168

RESULT 14

Q9MTA0 PRELIMINARY; PRT; 476 AA.

AC Q9MTA0; 09MTA1; 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

GN Probable intron maturase (Maturase K) (Fragment).

OS Cryptomeria japonica (Japanese cedar).

OC Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.

OX NCBI_TaxID=3369;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OB11, AND SAGA3; TISSUE=LEAF;

RA Kusumi J., Tsunura Y., Yoshimaru H., Tachida H.;

RT "Phylogenetic relationships in Taxodiaceae and Cupressaceae based on the matK, chlL, trnL-trnF IGS region and trnL intron sequences."

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II

CC INTRONS (BY SIMILARITY).

CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,

CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY

CC MITOCHONDRIAL INTRONS.

CC EMBL; AB030116; BAB01545.1; -.

DR InterPro: IPR000442; Intron_maturase2.

DR InterPro: IPR002866; MatK_N

DR Pfam; PF01348; Intron_maturase2; 1.

KW Chloroplast; mRNA processing.

FT NON TER 476

SO SEQUENCE 476 AA; 56791 MW; 7E7D491074C15PB2 CRC64;

QY Query Match 67.2%; Score 41; DB 8; Length 476;

Best Local Similarity 60.0%; Pred. No. 77;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VWRVRRWR 11

DB 159 LVRIFFRRWR 168

AC Q9M1S0; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CG9863 protein.
 GN CG9863.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX NCBI
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan H.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfenchko C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fodor C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glozier A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Jorgensen C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Rainer C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Strebbs R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhu D., Yang S., Yao Q.A.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RU EMBL; AF003460; AAF46983.1; -;
 DR EMBL; AF003460; AAF46983.1; -;
 DR FlyBase; FBgn003484; CG9863.
 DR InterPro; IPR001525; CS DNA_meth.
 DR PROSITE; PS00095; CS MTASE 2; UNKNOWN 1.
 SQ SEQUENCE 458 AA; 52193 MW; EB0482C86CDDCS CRC64;

Query Match 67.2%; Score 41; DB 5; Length 458;
 Best Local Similarity 87.5%; Pred. No. 74;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VYVYRRWV 10
 DB 91 VYVYRRWV 98

RESULT 10,
 ID Q9M1S2 PRELIMINARY; PRT; 462 AA.
 AC Q9M1S2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Probable intron maturase (Maturase K) (Fragment).
 GN MATK.
 OS Thuja standishii.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Coniferales; Cupressaceae; Thuja.
 NCBI_TaxID=89194;
 RX NCBI
 RP SEQUENCE FROM N.A.
 RC PubMed=10898782;
 RA Gadek P.A., Alpers D.L., Heslewood M.W., Quinn C.J.;
 RT "Relationships within Cupressaceae sensu lato: a combined morphological and molecular approach."
 RL Am. J. Bot. 87:1044-1057(2000).
 CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II INTRONS (BY SIMILARITY).
 CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS, BY AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY MITOCHONDRIAL INTRONS.
 DR EMBL; AF152215; AAF25768.1; -;
 DR InterPro; IPR000442; Intron_maturase2.
 DR InterPro; IPR002866; MatK_N.
 DR Pfam; PF01348; Intron_maturase2; 1.
 DR Pfam; PF01824; MatK_N; 1.
 KM Chloroplast; mRNA processing.
 FT NON_TER 1 1
 SQ SEQUENCE 462 AA; 55260 MW; F83B837AA2379HAD CRC64;

Query Match 67.2%; Score 41; DB 8; Length 462;
 Best Local Similarity 60.0%; Pred. No. 75;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VYVYRRWV 11
 DB 112 VYVYRRWV 121

RESULT 11
 ID Q9MSU4 PRELIMINARY; PRT; 470 AA.
 AC Q9MSU4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Probable intron maturase (Maturase K) (Fragment).
 GN MATK.
 OS Diselmia archeri.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Coniferales; Cupressaceae; Diselmia.
 NCBI_TaxID=13497;
 RX NCBI
 RP SEQUENCE FROM N.A.
 RC PubMed=10898782;
 RA Gadek P.A., Alpers D.L., Heslewood M.W., Quinn C.J.;
 RT "Relationships within Cupressaceae sensu lato: a combined morphological and molecular approach."
 RL Am. J. Bot. 87:1044-1057(2000).
 CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II INTRONS (BY SIMILARITY).
 CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS, BY AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY MITOCHONDRIAL INTRONS.
 DR EMBL; AF152193; AAF25746.1; -;
 DR InterPro; IPR000442; Intron_maturase2.
 DR InterPro; IPR002866; MatK_N.
 DR Pfam; PF01348; Intron_maturase2; 1.
 DR Pfam; PF01824; MatK_N; 1.
 KM Chloroplast; mRNA processing.
 FT NON_TER 1 1
 SQ SEQUENCE 470 470

```

CC      INTRONS (BY SIMILARITY).
CC      -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC      AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC      MITOCHONDRIAL INTRONS.
DR      EMBL: AF152175; AAF25728.2; -.
DR      InterPro: IPR000442; Intron_maturase2.
DR      InterPro: IPR002866; MatK_N.
DR      Pfam: PF01348; Intron_maturase2; 1.
DR      Pfam: PF01824; MatK_N; 1.
KW      Chloroplast; mRNA processing.
FT      NON_TER 510 510
SQ      SEQUENCE 510 AA; 61040 MW; F574F0957D39C201 CRC64;

QY      Query Match 68.9%; Score 42; DB 8; Length 510;
        Best Local Similarity 70.0%; Pred. No. 57;
        Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db      159 LVVFRMR 168

RESULT 6
Q9MST2 PRELIMINARY; PRT; 510 AA.
AC      Q9MST2;
DT      01-OCT-2000 (TReMBLrel. 15, Created)
DT      01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT      01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE      Probable intron maturase (Maturase K) (Fragment).
GN      MATK.
OS      Neocallitropsis araucarioidea.
OG      Chloroplast.
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
OC      Neocallitropsis.
OX      NCBI_Taxid=89193;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      PubMed=10898782;
RA      Gadek P.A., Alpers D.L., Heslewood M.M., Quinn C.J.;
RT      "Relationships within Cupressaceae sensu lato: a combined
RT      morphological and molecular approach.";
RL      Am. J. Bot. 87:1044-1057(2000).
CC      -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC      INTRONS (BY SIMILARITY).
CC      -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC      AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC      MITOCHONDRIAL INTRONS.
DR      EMBL: AF152205; AAF25758.1; -.
DR      InterPro: IPR000442; Intron_maturase2.
DR      InterPro: IPR002866; MatK_N.
DR      Pfam: PF01348; Intron_maturase2; 1.
DR      Pfam: PF01824; MatK_N; 1.
KW      Chloroplast; mRNA processing.
FT      NON_TER 510 510
SQ      SEQUENCE 510 AA; 60660 MW; D99F0ECA1874D02F CRC64;

QY      Query Match 68.9%; Score 42; DB 8; Length 510;
        Best Local Similarity 70.0%; Pred. No. 57;
        Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db      159 LVVFRMR 168

RESULT 7
Q8TX8 PRELIMINARY; PRT; 271 AA.
AC      Q8TX8;
DT      01-JUN-2002 (TReMBLrel. 21, Created)
DT      01-JUN-2002 (TReMBLrel. 21, Last sequence update)

```

```

DT      01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE      At2g616p.
OS      Drosophila melanogaster (fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_Taxid=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA      Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA      George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA      Miranda A., Mungall C.J., Nunoo U., Pacled U., Pargas V., Park S.,
RA      Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA      Celinker S.;
RL      Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AY089449; AAL90187.1; -.
SQ      SEQUENCE 271 AA; 31514 MW; 4D2082163A24B34E CRC64;

QY      Query Match 67.2%; Score 41; DB 5; Length 271;
        Best Local Similarity 87.5%; Pred. No. 45;
        Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db      145 VVVRMR 152

RESULT 8
Q97U20 PRELIMINARY; PRT; 286 AA.
AC      Q97U20;
DT      01-OCT-2001 (TReMBLrel. 18, Created)
DT      01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT      01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE      ABC transporter, permease (glucose).
GN      SSO2848.
OS      Sulfolobus solfataricus.
OC      Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC      Sulfolobus.
OX      NCBI_Taxid=2287;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 35092 / DSM 1617 / P2;
RX      MEDLINE=21332296; PubMed=11427726;
RA      She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA      Aways M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA      De Moers A., Etraus G., Fletcher C., Gordon F.M.K.,
RA      Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA      Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA      Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA      Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT      "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL      Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR      EMBL: AE006878; AAK42958.1; -.
DR      InterPro: IPR000515; BPD_transp.
DR      Pfam: PF00528; BPD_transp; 1.
DR      PROSITE: PS00402; BPD_TRANSF_INN_MEMBER_1.
KW      Complete proteome.
SQ      SEQUENCE 286 AA; 32285 MW; 2478218F158ECB59 CRC64;

QY      Query Match 67.2%; Score 41; DB 17; Length 286;
        Best Local Similarity 75.0%; Pred. No. 48;
        Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db      279 VVVRMR 286

RESULT 9
Q9W1S0 PRELIMINARY; PRT; 458 AA.
AC      Q9W1S0;

```

DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JUN-1998 (Tremblrel. 05, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Alpha-(1,3)-fucosyltransferase (EC 2.4.1.1).
GN FUCT.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11637;
RX MEDLINE=97407924; PubMed=9261148;
RA Martin S.L., Edbrooke M.R., Hodgman T.C., van den Eljnden D.H.,
RT Bird M.I.,
RT "Lewis X biosynthesis in Helicobacter pylori. Molecular cloning of an
RT alpha(1,3)-fucosyltransferase gene."
RT J. Biol. Chem. 272:21349-21356(1997).
DR EMBL; AF006039; AAB93985.1; -
DR InterPro: IPR001503; GT 10.
DR Pfam: PF00852; Glyco transf 10; 2.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 333 AA; 39154 MW; CFFCDA127E0A8C CRC64;

Query Match 70.5%; Score 43; DB 2; Length 333;
Best Local Similarity 54.5%; Pred. No. 26;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 2 VVRVRRVRR 12
Db 323 LRVRRVRRWK 333
:::|||||:::

RESULT 3
QYBUB6 PRELIMINARY; PRT; 391 AA.
AC QYBUB6;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical protein APE1602.
GN APE1602.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatake Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaei A., Kosugi H.,
RA Hoshoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Maeda S., Funahashi T., Tanaka T., Kudo Y.,
RA Yamazaki Y., Kuehida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RT DNA Res. 6:83-101(1999).
RL EMBL; AP000062; BAA80602.1; -
DR InterPro: IPR000719; BUK_dkinase.
DR InterPro: IPR000687; RIOI_UNK.
DR Pfam: PF01163; RIOI; 1.
DR SMART; SM00080; RIOI; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS01243; RIOI; 1.
KW ATP-binding; Hypothetical protein; Transferase; Complete proteome.
SQ SEQUENCE 391 AA; 44473 MW; 1E242B682EC26923 CRC64;

Query Match 68.9%; Score 42; DB 17; Length 391;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVRVRRVRRVRR 12
Db 108 RVRVRRVRRWK 119
|||:|:|:|:

RESULT 4
QYMSV7 PRELIMINARY; PRT; 485 AA.
ID QYMSV7;
AC QYMSV7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Probable intron maturase (Maturase K) (Fragment).
GN MATK.
OS Callitris rhomboides.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Cupressaceae; Callitris.
OX NCBI_TaxID=13383;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=10898782;
RA Gadek P.A., Alpers D.L., Heslewood M.M., Quinn C.J.;
RT "Relationships within Cupressaceae sensu lato: a combined
RT morphological and molecular approach."
RL Am. J. Bot. 87:1044-1057(2000).
CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC INTRONS (BY SIMILARITY).
CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS, BY
CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC MITOCHONDRIAL INTRONS.
DR EMBL; AF152180; AAF25733.2; -
DR InterPro: IPR000442; Intron_maturase2.
DR InterPro: IPR002866; MATK_N.
DR Pfam; PF01348; Intron_maturase2; 1.
DR Pfam; PF01824; MATK_N; 1.
KW Chloroplast; mRNA processing.
FT NON TER 1 485
FT NON TER 1 485
SQ SEQUENCE 485 AA; 57298 MW; 17C22EBF117789B3 CRC64;

Query Match 68.9%; Score 42; DB 8; Length 485;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 VVRVRRVRR 11
Db 134 LRVRRVRRWK 143
:::|||||:::

RESULT 5
QYMSW2 PRELIMINARY; PRT; 510 AA.
ID QYMSW2;
AC QYMSW2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Probable intron maturase (Maturase K) (Fragment).
GN MATK.
OS Actinostrobilus acuminatus.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Cupressaceae;
OC Actinostrobilus.
OX NCBI_TaxID=103962;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=10898782;
RA Gadek P.A., Alpers D.L., Heslewood M.M., Quinn C.J.;
RT "Relationships within Cupressaceae sensu lato: a combined
RT morphological and molecular approach."
RL Am. J. Bot. 87:1044-1057(2000).
CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II


```

RT human herpesvirus-6 containing homologues of human cytomegalovirus
RT major immediate-early and replication genes.";
RL Virology 204:738-750(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95266321; Pubmed=7747482;
RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
RA Martin M.E., Efstrathiou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
RT and genome evolution.";
RL Virology.209:29-51(1995).
CC -!- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL8,
CC EBV-1 54, VZV 52 AND HCMV 102.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U13194; AAA68465.1; -
CC DR EMBL; X83413; CAA58366.1; -
CC KM DNA replication.
SQ SEQUENCE 662 AA; 76316 MW; 2E29BBAF10BF7621 CRC64;
QY 3 VRYVRRRV 10
Db 625 LRYVRRRV 632

```

```

Query Match 59.0%; Score 36; DB 1; Length 662;
Best Local Similarity 75.0%; Pred. NO. 59;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Search completed: June 9, 2003, 11:56:40
Job time : 3.4253 secs

```

RA Rep1 A.M., Bresnick E.;
 RT "gene structure and nucleotide sequence for rat cytochrome P-450c."; Arch. Biochem. Biophys. 237:465-476(1985).
 RL [4]
 RN
 RP SUBCELLULAR LOCATION.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=96012193; PubMed=9348277;
 RA Addya S., Anandatheethavada H.K., Biswas G., Bhagwat S.V., Mullick J., Avadhani N.G.;
 RT "Targeting of NH2-terminal-processed microsomal protein to mitochondria: a novel pathway for the biogenesis of hepatic mitochondrial P450MT2.";
 RL J. Cell Biol. 139:589-599(1997).
 CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND; ENDOPLASMIC RETICULUM AND MITOCHONDRIAL.
 CC -1- TISSUE SPECIFICITY: LIVER.
 CC -1- INDUCTION: BY 3-METHYLCOLANTHRENE (3MC) AND BETA-NAPHTHOLAVONE (BNP).
 CC -1- DOMAIN: CONTAINS A CHIMERIC SIGNAL THAT FACILITATES TARGETING OF THE PROTEIN TO BOTH THE ENDOPLASMIC RETICULUM AND MITOCHONDRIA. A 12 AMINO ACID SEQUENCE BETWEEN 33 AND 44 FUNCTIONS AS A PUTATIVE MITOCHONDRIAL-TARGETING SIGNAL. THE REMOVAL OF THE FIRST 4-OR 32- AMINO ACID RESIDUES FROM THE INACT PROTEIN POSITIONS THE MITOCHONDRIAL TARGETING SIGNAL FOR EFFICIENT BINDING TO THE MITOCHONDRIAL IMPORT RECEPTORS. THE MEMBRANE-FREE P450I1 SEEMS TO BE MORE SENSIBLE TO PROTEOLYSIS.
 CC -1- PTM: TWO FORMS, MT2A (LONG FORM) AND MT2B (SHORT FORM); ARE PRODUCED BY NH2-TERMINAL PROTEOLYTIC CLEAVAGE. THIS CLEAVAGE ACTIVATES A CRYPTIC MITOCHONDRIAL TARGETING SIGNAL.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: K03246; AAA1027.1; -;
 DR EMBL: X00469; CA25153.1; -;
 DR EMBL: M26129; AAA1025.1; -;
 DR PIR: A00185; O4RTWC.
 DR PIR: S45716; S45716.
 DR HSSP: P00179; 1D76.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 DR KMW Oxioreductase; Monooxygenase; Electron transport; Membrane; Heme; Microsome; Endoplasmic reticulum; Mitochondrion.
 FT CHAIN 1 524 CYTOCHROME P450 IAI.
 FT PROPEP 1 524
 FT CHAIN 5 524 CYTOCHROME P450MT2A.
 FT PROPEP 1 32
 FT CHAIN 33 524 CYTOCHROME P450MT2B.
 FT BINDING 461 461 HEME.
 FT DOMAIN 33 44 MITOCHONDRIAL-TARGETING SIGNAL.
 FT MUTAGEN 33 33 VT->AI: NO PROTEOLYTIC CLEAVAGE.
 FT CONFLICT 53 53 I -> M (IN REF. 2).
 FT CONFLICT 494 494 M -> S (IN REF. 3).
 SQ SEQUENCE 524 AA; 59393 MW; C766DF8044D598C5 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 524;
 Best Local Similarity 77.8%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRVRRVR 10
 DB 29 VVRVRRVR 37
 RESULT 14
 ID TR2A_PSEFL STANDARD; PRT; 558 AA.
 AC Q51761;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Putative transposase for insertion sequence IS1162.
 OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 CC NCBI_TaxID=294;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ST;
 RX MEDLINE=95212933; PubMed=7698671;
 RA Solinas F., Marconi A.M., Ruzzi M., Zennaro B.;
 RT "Characterization and sequence of a novel insertion sequence, IS1162, from Pseudomonas fluorescens.";
 RL Gene 155:77-82(1995).
 CC -1- FUNCTION: REQUIRED FOR THE TRANSPOSITION OF THE INSERTION ELEMENT (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE IS2/IS408/IS1162 FAMILY OF TRANSPOSASES.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X79443; CA55959.1; -;
 DR InterPro: IPR001584; Rve.
 DR Pfam: PF00665; Rve; 1.
 DR Transposable element; Transposition; DNA-binding; DNA recombination.
 FT DNA BIND 25 44 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 558 AA; 62990 MW; D46250C5071F709E CRC64;

Query Match 59.0%; Score 36; DB 1; Length 558;
 Best Local Similarity 60.0%; Pred. No. 49;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVRVRRVR 12
 DB 265 VVRVRRVR 274
 RESULT 15
 ID HEPA_HSV6U STANDARD; PRT; 662 AA.
 AC P52375;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 GN U74 OR HDR1.
 OS Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Roseolovirinae.
 CC NCBI_TaxID=10370;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95027704; PubMed=7941342;
 RA Nicholas J.;
 RT "Nucleotide sequence analysis of a 21-kbp region of the genome of

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL, X59668; CAA42201.1; ALT INIT.
DR InterPro; IPR000636; M+channel_nlg.
DR InterPro; IPR000595; CNMP-binding.
DR Pfam; PF00027; CNMP-binding; 1.
DR Pfam; PF00520; Ion Trans; 1.
DR SMART; SM00100; CNMP; 1.
DR PROSITE; PS00886; CNMP-BINDING_1; 1.
DR PROSITE; PS00889; CNMP-BINDING_2; 1.
DR PROSITE; PS50042; CNMP-BINDING_3; 1.
DR Ionic channel; Ion transport; CAMP-binding; Transmembrane;
KM Multigene family; Olfaction.
FT DOMAIN 1 140 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 141 160 H1 (POTENTIAL).
FT DOMAIN 161 173 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 174 192 H2 (POTENTIAL).
FT DOMAIN 193 216 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 217 236 H3 (POTENTIAL).
FT DOMAIN 237 274 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 275 297 H4 (POTENTIAL).
FT DOMAIN 298 349 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 350 369 H5 (POTENTIAL).
FT DOMAIN 370 453 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 454 474 H6 (POTENTIAL).
FT DOMAIN 475 664 CYTOPLASMIC (POTENTIAL).
FT NP BIND 462 584 CAMP (BY SIMILARITY).
FT BINDING 521 521 CAMP (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 664 AA; 76205 MW; SE9170DOB322B3B9 CRC64;

Query Match 60.7%; Score 37; DB 1; Length 664;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RVVRVVRVVR 12
Db 71 RLWGVIRQMANR 82

RESULT 12

TRY3 ANOGA STANDARD; PRT; 275 AA.

AC P35037;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin 3 precursor (EC 3.4.21.4).

OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Culicidae; Anopheles.
OX NCBI_TaxID=7165;

RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Suakoko; TISSUE=Midgut;
RX MEDLINE=93327779; PubMed=8335004;
RA Mueller H.M., Crampson J.W., della Torre A., Sindén R., Crisanti A.;
RT "Members of a trypsin gene family in Anopheles gambiae are induced in
the gut by blood meal."
RL EMBL J. 12:2891-2900(1993).

CC -1- FUNCTION: MAJOR FUNCTION MAY BE TO AID IN DIGESTION OF THE BLOOD
CC MEAL.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MIDGUT.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL, Z22930; CAA80517.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.112; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP SPc; 1.
DR PROSITE; PS50240; TRYP SIN DOM; 1.
DR PROSITE; PS00134; TRYP SIN_HIS; 1.
DR PROSITE; PS00135; TRYP SIN_SER; 1.
KM Hydrolyase; Serine protease; Signal; Zymogen; Digestion;
KM Multigene family.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 48 ACTIVATION PEPTIDE.
FT CHAIN 49 275 TRYP SIN 3.
FT ACT_SITE 89 89 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 134 134 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 230 230 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 74 90 BY SIMILARITY.
FT DISULFID 199 215 BY SIMILARITY.
FT DISULFID 226 250 BY SIMILARITY.
FT SITE 224 224 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 275 AA; 29403 MW; 2APFP2D507FEE89F7 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 275;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 RVVRVVRVVR 11
Db 262 VAVVRNMR 270

RESULT 13

CP11 RAT STANDARD; PRT; 524 AA.

AC P00185;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome P450 1A1 (EC 1.14.14.1) (CYP1A1) (P450-C)
DE (P450MT2).

GN CYP1A1 OR CYP1A-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

RA [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8428082; PubMed=6089174;
RA Sogawa K., Gotch O., Kawajiri K., Fujii-Kuriyama Y.;
RT "Distinct organization of methylcholanthrene- and phenobarbital-
induced cytochrome P-450 genes in the rat."
RL Proc. Natl. Acad. Sci. U.S.A. 81:5066-5070(1984).

CC -1- FUNCTION: MAJOR FUNCTION MAY BE TO AID IN DIGESTION OF THE BLOOD
CC MEAL.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MIDGUT.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

OS pinda contorta (Shore pine) (Lodgepole pine).
 CC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 CC NCBI_TaxID=3339;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=91279438; PubMed=2057349;
 RA Lidholm J.A., Gustafsson P.;
 RT "A three-step model for the rearrangement of the chloroplast
 RT trnK-psbA region of the gymnosperm Pinus contorta.";
 RL Nucleic Acids Res. 19:2881-2887(1991).
 CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
 CC INTRONS.
 CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
 CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED
 CC BY MITOCHONDRIAL INTRONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X57097; CAA40382.1; -.
 CC DR PIR; S22550; S22550.
 CC DR InterPro; IPR000442; Intron_mature2.
 CC DR InterPro; IPR002866; MatK_N.
 CC DR Pfam; PF01348; Intron_mature2; 1.
 CC DR Pfam; PF01824; MatK_N; 1.
 CC KM Chloroplast; mRNA processing.
 CC SO SEQUENCE 515 AA; 60872 MW; 3EF62652D7CF39 CRC64;
 Query Match 62.3%; Score 38; DB 1; Length 515;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 VVRVRRVVR 11
 DB 160 LVRTFRWR 169
 RESULT 10
 MATK_PINTH STANDARD; PRT; 515 AA.
 AC Q00866;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable intron maturase (Maturase K).
 GN MATK OR YCF14.
 OS Pinus thunbergii (Green pine) (Japanese black pine).
 CC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 CC NCBI_TaxID=3350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92212283; PubMed=1557027;
 RA Tsudzuki T., Nakashima K., Tsudzuki T., Hiraoka J., Shibata M.,
 RA Wakaugui T., Sugiura M.;
 RT "Chloroplast DNA of black pine retains a residual inverted repeat
 RT lacking rRNA genes; nucleotide sequences of trnQ, trnK, psbA, trnI
 RT and trnH and the absence of trnS6.";
 RL Mol. Gen. Genet. 232:206-214(1992).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=95024047; PubMed=7937893;
 CC Wakaugui T., Tsudzuki T., Ito S., Nakashima K., Tsudzuki T.,
 CC Sugiura M.;

RT "Loss of all rnh genes as determined by sequencing the entire
 RT chloroplast genome of the black pine Pinus thunbergii.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
 CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
 CC INTRONS.
 CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
 CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED
 CC BY MITOCHONDRIAL INTRONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; D11467; BAA02022.1; -.
 CC DR PIR; D17510; BAA04308.1; -.
 CC DR PIR; S20449; S20449.
 CC DR InterPro; IPR000442; Intron_mature2.
 CC DR InterPro; IPR002866; MatK_N.
 CC DR Pfam; PF01348; Intron_mature2; 1.
 CC DR Pfam; PF01824; MatK_N; 1.
 CC KM Chloroplast; mRNA processing.
 CC SO SEQUENCE 515 AA; 60794 MW; A0CA483583EB87D2 CRC64;
 Query Match 62.3%; Score 38; DB 1; Length 515;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 VVRVRRVVR 11
 DB 160 LVRTFRWR 169
 RESULT 11
 CNG2_RABIT STANDARD; PRT; 664 AA.
 AC Q28718;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cyclic-nucleotide-gated olfactory channel (Cyclic-nucleotide-gated
 DE cation channel 2) (CNG channel 2) (CNG2) (Aorta CNG channel)
 DE (RACNG).
 GN CNG2 OR CNG2.
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aorta;
 RX MEDLINE=93359035; PubMed=7689061;
 RA Biel M., Altenhofen W., Hüllin R., Ludwig J., Freichel M.,
 RA Flockerzi V., Dascal N., Kaupp U.B., Hofmann F.;
 RT "Primary structure and functional expression of a cyclic nucleotide-
 RT gated channel from rabbit aorta.";
 RL FEBS Lett. 329:134-138(1993).
 CC -1- FUNCTION: ODORANT SIGNAL TRANSDUCTION IS PROBABLY MEDIATED
 CC BY A G-PROTEIN COUPLED CASCADE USING CAMP AS SECOND MESSENGER.
 CC THE OLFACTORIC CHANNEL CAN BE SHOWN TO BE ACTIVATED BY CYCLIC
 CC NUCLEOTIDES WHICH LEADS TO A DEPOLARIZATION OF OLFACTORIC
 CC SENSORY NEURONS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

QY 1 RVRVVRVVR 12
: ||| |||
Db 257 KVASVARRP 268

RESULT 7

NUOF MYCTU STANDARD; PRT; 445 AA.
AC P95176;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH dehydrogenase I chain F (EC 1.6.5.3) (NADH-ubiquinone
oxidoreductase chain F).
GN NUOF OR RV3150 OR MT3238 OR MTCY03A2.08C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Saltsion J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleishmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayan L.A., Esmolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBD databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- COFACTOR: FMN AND ONE 4FE-4S CLUSTER (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 51 KDa SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z83867; CAB06289.1; -.
DR EMBL; AE007138; AAK47577.1; -.
DR TIGR; MT3238; -.
DR TubercuList; RV3150; -.
DR InterPro; IPR001949; Complex1_51K.
DR Pfam; PF01512; Complex1_51K; 1.
DR PROSITE; PS00644; Complex1_51K; 1.
DR PROSITE; PS00645; Complex1_51K; 2; 1.
KW Oxidoreductase; NAD; Ubiquinone; Flavoprotein; FMN; Iron-sulfur;
KW 4Fe-4S; Complete proteome.
KM NP_BIND 61 70 NAD (BY SIMILARITY).
FT NP_BIND 177 224 FMN (BY SIMILARITY).
FT METAL 353 353 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 356 356 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 359 359 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 399 399 IRON-SULFUR (4FE-4S) (POTENTIAL).

SEQ SEQUENCE 445 AA; 48101 MW; C02A660F43842926 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 445;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVVRVVR 9
: ||| |||
Db 337 VVRVVRVVR 344

RESULT 8

MARK TORCL STANDARD; PRT; 495 AA.
AC 09MV77;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable intron maturase (Maturase K).
GN MARK.
OS Torreyia californica (California nutmeg).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Torreya.
OX NCBI_TaxID=89482;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=20179521; PubMed=10712841;
RA Cheng Y., Nicolson R.G., Tripp K., Chaw S.,
RT "Phylogeny of taxaceae and Cephalotaxaceae genera inferred from
RT chloroplast mark gene and nuclear rDNA ITS region.";
RL Mol. Phylogenet. Evol. 14:353-365(2000).
CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC INTRONS.
CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED
CC BY MITOCHONDRIAL INTRONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB023998; BAAB6047.1; -.
DR InterPro; IPR000442; Intron_maturase2.
DR Pfam; PF01348; Intron_maturase2; 1.
DR Pfam; PF01824; Mark_N1.
KW Chloroplast; mRNA processing.
KW SEQUENCE 495 AA; 59260 MW; 939PB037285A36F8 CRC64;
Query Match 62.3%; Score 38; DB 1; Length 495;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRVVRVVR 11
: ||| |||
Db 159 LVTRPRMR 168

RESULT 9

MARK PINCO STANDARD; PRT; 515 AA.
AC P24685;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable intron maturase (Maturase K).
GN MARK OR YCF14.

RT "Determinants of site-specific recombination in the lambdoid
coliphage HK022. An evolutionary change in specificity."
RL J. Mol. Biol. 207:695-717(1998).
CC -1- FUNCTION: EXCISIONASE AND INTEGRASE ARE NECESSARY FOR THE EXCISION
OF PROPHAGE FROM THE HOST GENOME BY SITE-SPECIFIC RECOMBINATION AT
THE ATT SITE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC
DR EMBL; M60848; AAA67901.1; -
DR EMBL; X51962; CAA36222.1; -
DR PIR; S06533; S06533.
KW DNA recombination; DNA-binding.
SQ SEQUENCE 72 AA; 8635 MW; 0E6A4843503344AA CRC64;

QY Query Match 62.3%; Score 38; DB 1; Length 72;
Best Local Similarity 63.6%; Pred. No. 2.3;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 1 RVRVVRVRVR 11
16 RSLVTVRRVR 26

RESULT 5
VXIS_LAMB
ID VXIS_LAMB STANDARD; PRT; 72 AA.
AC P03639;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Excisionase.
GN XIS.
OS Bacteriophage lambda.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10710;
RN [1]
RX MEDLINE=63189071; PubMed=6221115;
RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;
RT "Nucleotide sequence of bacteriophage lambda DNA."
RL J. Mol. Biol. 162:729-773(1982).
RN [2]
RX MEDLINE=61053845; PubMed=6253947;
RA Davies R.W.;
RT "DNA sequence of the int-xis-Pi region of the bacteriophage lambda;
RT overlap of the int and xis genes."
RL Nucleic Acids Res. 8:1765-1782(1980).
RN [3]
RX MEDLINE=80234646; PubMed=6446713;
RA Hoess R.H., Foeller C., Bidwell K., Landy A.;
RT "Site-specific recombination functions of bacteriophage lambda: DNA
RT sequence of regulatory regions and overlapping structural genes for
RT int and xis."
RL Proc. Natl. Acad. Sci. U.S.A. 77:2482-2486(1980).
CC -1- FUNCTION: EXCISIONASE AND INTEGRASE ARE NECESSARY FOR THE EXCISION
OF PROPHAGE FROM THE HOST GENOME BY SITE-SPECIFIC RECOMBINATION AT
THE ATT SITE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC
DR EMBL; J02459; AAA96563.1; -
DR PIR; A04321; RSBPXL.
KW DNA recombination; DNA-binding.
SQ SEQUENCE 72 AA; 8605 MW; 0E6A4843502200AA CRC64;

QY Query Match 62.3%; Score 38; DB 1; Length 72;
Best Local Similarity 63.6%; Pred. No. 2.3;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 1 RVRVVRVRVR 11
16 RSLVTVRRVR 26

RESULT 6
AMPC_SERMA
ID AMPC_SERMA STANDARD; PRT; 376 AA.
AC P18539;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Beta-lactamase precursor (EC 3.5.2.6) (Cephalosporinase).
GN AMPc.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615;
RN [1]
RX MEDLINE=91032898; PubMed=2227364;
RA Nomura K., Yoshida T.;
RT "Nucleotide sequence of the Serratia marcescens SR50 chromosomal ampc
RT beta-lactamase gene."
RL FEBS Microbiol. Lett. 58:295-299(1990).
CC -1- FUNCTION: THIS PROTEIN IS A SERINE BETA-LACTAMASE WITH A SUBSTRATE
CC SPECIFICITY FOR CEPHALOSPORINS.
CC -1- CATALYTIC ACTIVITY: A beta-lactam + H(2)O = a substituted beta-
CC amino acid.
CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE CLASS-C BETA-LACTAMASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC
DR EMBL; X52964; CAA37137.1; -
DR PIR; S11710; QKSE.
DR PIR; A48176; A48176.
DR HSSP; P05364; 2BLT.
DR MEROPS; S12.UMN; -
DR InterPro; IPR001466; Beta_lactamase.
DR InterPro; IPR001586; Beta_lactamase_C.
DR Pfam; PF00144; beta_lactamase; 1.
DR PROSITE; PS00336; BETA_LACTAMASE_C; 1.
KW Hydrolase; Antibiotic resistance; Periplasmic; Signal.
FT SIGNAL 1 21
FT CHAIN 22 376 BETA-LACTAMASE.
FT ACT SITE 79 79 BY SIMILARITY.
FT BINDING 328 330 SUBSTRATE (BY SIMILARITY);
SQ SEQUENCE 376 AA; 41096 MW; 890CEEC7C27925150 CRC64;

QY Query Match 62.3%; Score 38; DB 1; Length 376;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

ID MATK_CUNLA STANDARD; PRT; 508 AA.
 AC Q9MSV2; Q9MT93;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Probable intron maturase (Maturase K).
 OS Cunninghamia lanceolata (China fir).
 OC Chloroplast.
 OC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
 OC Cunninghamia.
 NCBI_TaxID=28977;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed:10898782;
 RA Gadek P.A., Alpers D.L., Heslewood M.M., Quinn C.J.;
 RT "Relationships within Cupressaceae sensu lato: a combined
 morphological and molecular approach.";
 RL Am. J. Bot. 87:1044-1057(2000).
 RN [2]
 RP SEQUENCE OF 1-477 FROM N.A.
 RC TISSUE=Leaf;
 RA Kusumi J., Tsunuma Y., Yoshimaru H., Tachida H.;
 RT "Phylogenetic relationships in Taxodiaceae and Cupressaceae based on
 the matk, chl, trnL-trnF, rbcL and trnL intron sequences.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
 INTRONS.
 CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
 AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED
 BY MITOCHONDRIAL INTRONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF152185; AAF25738.1; -;
 DR EMBL; AB030125; BAB01594.1; -;
 DR InterPro: IPR000442; Intron_maturase2.
 DR InterPro: IPR002866; Matk_N.
 DR Pfam; PF01348; Intron_maturase2; 2.
 DR Pfam; PF01824; Matk_N; 1.
 KM Chloroplast; mRNA processing.
 SQ SEQUENCE 508 AA; 60732 MW; 9003C19B127B562 CRC64;
 Query Match 67.2%; Score 41; DB 1; Length 508;
 Best Local Similarity 60.0%; Pred. No. 5.8;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VVRVVRVVR 11
 DB 159 LVRIFFRMIR 168
 RESULT 3
 ID RL31_AERPE STANDARD; PRT; 105 AA.
 AC Q9YD5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 50S ribosomal protein L31e.
 GN RPL31E OR APE1087.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
 OC Desulfurococcaceae; Aeropyrum.
 NCBI_TaxID=56636;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatake Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Koeugi H.,
 RA Hosoyama A., Fukui S., Nagai S., Nishijima K., Nakazawa H.,
 RA Takamitsu M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 CC -1- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AP00060; BAA80072.1; -;
 DR InterPro: IPR00054; Ribosomal_L31e.
 DR Pfam; PF0198; Ribosomal_L31e; 1.
 DR PROSITE; PS01144; RIBOSOMAL_L31E; FALSE_NEG.
 KM Ribosomal protein; Complete proteome.
 SQ SEQUENCE 105 AA; 12527 MW; 7E5DF7999E74A098 CRC64;
 Query Match 63.9%; Score 39; DB 1; Length 105;
 Best Local Similarity 66.7%; Pred. No. 2.4;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RVRVVRVVR 12
 DB 28 RAVRVVRVVR 39
 RESULT 4
 ID VXS_BP434 STANDARD; PRT; 72 AA.
 AC P11683; P16408;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Excisionase.
 GN Xis.
 OS Bacteriophage 434, and
 OS Bacteriophage HK022.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 OC Lambda-like viruses.
 NCBI_TaxID=10712, 10742;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Phage 434;
 RX MEDLINE=9134614; PubMed=1715186;
 RA Baker J., Limberger R., Schneider S.J., Campbell A.;
 RT "Recombination and modular exchange in the genomes of new lambdaoid
 phages.";
 RL New Biol. 3:297-308(1991).
 RN [2]
 RP SEQUENCE OF 1-64 FROM N.A.
 RC SPECIES=Phage 434;
 RX MEDLINE=88167849; PubMed=2965063;
 RA Limberger R.J., Campbell A.W.;
 RT "Functional elements of DNA upstream from the integrase operon that
 are conserved in bacteriophages 434 and lambda.";
 RL Gene 61:135-144(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Phage HK022;
 RX MEDLINE=89342457; PubMed=2547971;
 RA Vagli E., Dolev S., Oberio J., Kislav N., Ramalho N., Weisberg R.A.;

RESULT 12

Probable permease of ABC transporter PA5155 (imported) - *Pseudomonas aeruginosa* (strain C:\Species: *Pseudomonas aeruginosa*
C:\Date: 15-Sep-2000 #sequence revision 15-Sep-2000 #text change 31-Dec-2000
C:\Accession: H83001
C:\Stover, C.K., Plam, X.Q.; Ewain, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:\Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:\Reference number: AB2950; MUID:20437337; PMID:10584043
A:\Accession: H83001
A:\Status: Preliminary
A:\Molecule type: DNA
A:\Residues: 1-230 <STO>
A:\Cross-references: GB:AE004928; GB:AE004091; NID:99951450; PIDN:AG08540.1; GSPDB:GN001001
A:\Experimental source: strain PA01
C:\Genetics:
A:\Gene: PA5155
C:\Superfamily: histidine permease protein M

RESULT 13

beta-lactamase (EC 3.5.2.6) precursor - *Serratia marcescens*
C.Species: *Serratia marcescens*
C.Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Jun-1999
C.Accession: A48176, S11710
R.Nomura, K.; Yoshida, T.
FEMS Microbiol. Lett. 70, 295-300, 1990
A.Title: Nucleotide sequence of the *Serratia marcescens* SRS0 chromosomal ampc beta-lacta
A.Reference number: A48176
A.Accession: A48176
A.Molecule type: DNA
A.Residues: 1-376 <NM>
A.Cross-references: EMBL:X52964, NID:947223, PIDN:CA43137.1, PID:947224
A.Note: Submitted to the EMBL Data Library, May 1990
A.Note: the authors translated the codon GTA for residue 258 as Asp
C.Genetics:
A.Gene: ampc
C.Superfamily: *Escherichia coli* beta-lactamase
C.Keywords: antibiotic resistance, hydrolase
P.1-31/Domain: signal sequence #status predicted <SIG>
P.122-376/Product: beta-lactamase #status predicted <M>
F.72-376/Product: beta-lactamase #status predicted <M>
P.79/Active site: Ser #status predicted

RESULT 14

NM22 dehydrogenase (ubiquinone) [EC 1.6.5.3] I chain F - Mycobacterium tuberculosis (strain H37Rv)
 Cspecies: Mycobacterium tuberculosis
 Cdate: 14-May-1999 #sequence_revision 14-May-1999 #text_change 03-Jun-2002
 Caccession: G1064
 R:COLE, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.

Nature 393, 537-544, 1998

A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70647
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-445 <COI>
A:Cross-references: GB:Z63867; GB:AL123456; NID:93261695; PIDN:CA806289.1; PID:g1781218
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: nuoF
C:Function:
A:Description: catalyzes the reduction of ubiquinone to ubiquinol by NADH
C:Pathway: oxidative phosphorylation
A:Superfamily: NADH dehydrogenase (ubiquinone) chain F; NADH dehydrogenase (ubiquinone)
C:Keywords: 4Fe-4S; flavoprotein; FMN; iron-sulfur protein; metalloprotein; NAD; oxidat
F:43-421/Domain: NADH dehydrogenase (ubiquinone) I chain F homology <NUOF>
F:61-70/Region: NAD binding motif
F:177-192/Region: FMN binding motif
F:355_356_359/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

RESULT 15

probable maturase, intron encoded - Japanese black pine chloroplast
N:Alternate names: hypothetical protein 515, crnk 3 -region
C:Species: chloroplast Pinus thunbergiana (Japanese black pine)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C:Accession: S20449; T07428
R:Tsudzuki, J.; Nakashima, K.; Tsudzuki, T.; Hiratsuka, J.; Shibata, M.; Wakasugi, T.;
Mol. Gen. Genet. 232, 206-214, 1992
A:title: Chloroplast DNA of black pine retains a residual inverted repeat lacking rRNA
A:Reference number: S20449; PMID:92212283; PMID:1557027
A:Accession: S20449
A:Molecule type: DNA
A:Residues: 1-515 <TSU>
A:Cross-references: EMBL:D11467; NID:G344007; PIDD:BA02022.1; PID:G344010
R:Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugita, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A:title: Loss of all nch genes as determined by sequencing the entire chloroplast genome
A:Reference number: z16030; PMID:95024047; PMID:7937893
A:Accession: T07428
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-515 <WAK>
A:Cross-references: EMBL:D17510; NID:G529643; PIDD:BA04308.1; PID:G1262589
C:Genetics:
A:Gene: matK
A:Genome: chloroplast
C:Superfamily: barley chloroplast maturase matK
C:Keywords: chloroplast

Matches: 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVVVVRRRVR 12

DB 466 LKRTIRRWVK 476

RESULT 8
excisionase - phage lambda
C/Species: phage lambda

C/Date: 31-Oct-1980 #sequence_revision 23-Oct-1981 #text_change 23-Jul-1999

C/Accession: C94164; C43012; A93699; A93844; A04321

R/Panels, D. submitted to the Nucleic Acid Sequence Database, September 1982

A/Reference number: A94614

A/Accession: C94164

A/Molecule type: DNA

A/Residues: 1-72 <DNA>

R/Sanger, F.; Coulson, A.R.; Hong, G.F.; Hill, D.F.; Petersen, G.B.

J. Mol. Biol. 162, 729-773, 1982

A/Title: Nucleotide sequence of bacteriophage lambda DNA.

A/Reference number: A92891; MUID:83189071; PMID:6221115

A/Accession: C43012

A/Molecule type: DNA

A/Residues: 1-72 <DNA>

A/Cross-references: GB:J02459; GB:M17233; GB:M24325; GB:V00636; GB:X00906; NID:G215104;

R/Davies, R.W.

Nucleic Acids Res. 8, 1765-1782, 1980

A/Title: DNA sequence of the int-xis P-1 region of the bacteriophage lambda; overlap of

A/Reference number: A93699; MUID:81053845; PMID:6253947

A/Accession: A93699

A/Molecule type: DNA

A/Residues: 1-72 <DNA>

A/Cross-references: GB:J02459; GB:M17233; GB:M24325; GB:V00636; GB:X00906; NID:G215104;

R/Hoess, R.H.; Feller, C.; Bidwell, K.; Landy, A.

Proc. Natl. Acad. Sci. U.S.A. 77, 2482-2486, 1980

A/Title: Site-specific recombination functions of bacteriophage lambda: DNA sequence of

A/Reference number: A93844; MUID:80234646; PMID:6446713

A/Accession: A93844

A/Molecule type: DNA

A/Residues: 1-72 <HOB>

A/Cross-references: GB:J02459; GB:M17233; GB:M24325; GB:V00636; GB:X00906; NID:G215104;

C/Comment: Excisionase and lambda integrase are necessary for the excision of prophage 1

C/Genetics: xis

A/Gene: xis

A/Map position: 59,95-59,51

C/Superfamily: phage lambda excisionase

C/Keywords: DNA binding

Query Match

Best Local Similarity 62.3%; Score 38; DB 1; Length 72;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVVVRVRVR 11

DB 16 RSLRTVRVR 26

RESULT 9

excisionase - phage HK022

C/Species: phage HK022

C/Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999

C/Accession: S06533

R/Yagil, E.; Dolev, S.; Oberto, J.; Kislav, N.; Ramaiha, N.; Weisberg, R.A.

J. Mol. Biol. 207, 695-717, 1989

A/Title: Determinants of site-specific recombination in the lambdaoid coliphage HK022. An

A/Reference number: S04990; MUID:89342457; PMID:2547971

A/Accession: S06533

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-72 <YAG>

A/Cross-references: EMBL:X51962; NID:G15759; PIDD:CAA36222.1; PIDD:G15761

C/Genetics: xis

A/Gene: xis

C/Superfamily: phage lambda excisionase

C/Keywords: DNA binding

Query Match

Best Local Similarity 62.3%; Score 38; DB 2; Length 72;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVVVRVRVR 11

DB 16 RSLRTVRVR 26

RESULT 10

excisionase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C/Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002

C/Accession: A90729

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A/Reference number: A99629; MUID:21156231; PMID:11258796

A/Accession: A90729

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-72 <HAY>

A/Cross-references: GB:BA000007; PIDD:BA034224.1; PIDD:G13360260; GSPDB:GNO0154

A/Experimental source: strain O157:H7, substrain RIMD 0509952

C/Genetics: xis

A/Gene: ECE0801

C/Superfamily: phage lambda excisionase

Query Match

Best Local Similarity 62.3%; Score 38; DB 2; Length 72;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVVVRVRVR 11

DB 16 RSLRTVRVR 26

RESULT 11

hypothetical protein 1313 - Sphingomonas aromaticivorans plasmid pLNT

C/Species: Sphingomonas aromaticivorans

C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000

C/Accession: T31302

R/Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.;

submitted to the EMBL Data Library, July 1998

A/Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati

A/Reference number: Z20992

A/Accession: T31302

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-201 <ROM>

A/Cross-references: EMBL:AF079317; NID:G3378261; PIDD:G3378443; PIDD:AA04026.1

C/Genetics: xis

A/Gene: plasmid pLNT1

A/Note: orf1313

Query Match

Best Local Similarity 62.3%; Score 38; DB 2; Length 201;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVVVRVRVR 12

DB 110 VRAVVRVRVR 119

OY 5 VVRVRRVVR 12
|:|:|:|:
Db 279 VVRVRRVVR 286

RESULT 3

T46005
receptor-like protein kinase - Arabidopsis thaliana

N:Alternate names: protein T10K17.40

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Mar-2000

C:Accession: T46005

R:Benes, V.; Wurmbach, R.; Drzonek, H.; Ansozge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23019

A:Accession: T46005

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-662 <BEN>

A:Cross-references: EMBL:AL132977

A:Experimental source: cultivar Columbia; BAC clone T10K17

C:Genetics:

A:Map position: 3

A:Insertions: 447/1

A:Note: T10K17.40

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 65.6%; Score 40; DB 2; Length 662;

Best Local Similarity 70.0%; Pred. No. 56;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VVRVRRVVR 11
|:|:|:|:
Db 596 LVRRVRRVVR 605

RESULT 4

H72708
probable ribosomal protein L31 APE1087 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 02-Aug-2002

C:Accession: H72708

R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takat

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K

DNA Res. 6: 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: H72708

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-105 <KAW>

A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BA80072.1; PID:dl043856; PID:G510

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1087

C:Superfamily: rat ribosomal protein L31

Query Match 63.9%; Score 39; DB 2; Length 105;

Best Local Similarity 66.7%; Pred. No. 16;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 RVVRVRRVVR 12
|:|:|:|:
Db 28 RAVRVRRVVR 39

RESULT 5

G71862
alpha-(1,3)-fucosyltransferase - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999

C:Accession: G71862

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: G71862

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-436 <ARN>

A:Cross-references: GB:AE001528; GB:AE001439; NID:G4155575; PIDN:AAD06573.1; PID:G41555

A:Experimental source: strain J99

Query Match 63.9%; Score 39; DB 2; Length 436;

Best Local Similarity 45.5%; Pred. No. 56;

Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 VVRVRRVVR 12
|:|:|:|:
Db 423 LVRRVRRVVR 433

RESULT 6

B71914
alpha (1,3)-fucosyltransferase - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999

C:Accession: B71914

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: B71914

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-454 <ARN>

A:Cross-references: GB:AE001491; GB:AE001439; NID:G4155127; PIDN:AAD06169.1; PID:G41551

A:Experimental source: strain J99

C:Genetics:

A:Gene: fucT

Query Match 63.9%; Score 39; DB 2; Length 454;

Best Local Similarity 45.5%; Pred. No. 58;

Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 VVRVRRVVR 12
|:|:|:|:
Db 444 LVRRVRRVVR 454

RESULT 7

G64601
fucosyltransferase - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: G64601

R:Tom, J.F.; White, O.; Karlvage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalek, H.G.; Glodek, A.; McKenn

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Bordovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: G64601

A>Status: preliminary;

A:Molecule type: nucleic acid sequence not shown; translation not shown

A:Residues: 1-476 <TCM>

A:Cross-references: GB:AE000578; GB:AE000511; NID:G2313759; PIDN:AAD07710.1; PID:G23137

Query Match 63.9%; Score 39; DB 2; Length 476;

Best Local Similarity 45.5%; Pred. No. 60;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:48:06 ; Search time 4.97872 Seconds
(without alignments)
231.709 Million cell updates/sec

Title: US-10-079-075-9

Perfect score: 61

Sequence: 1 RVRRVVRWVR 12

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 73: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	68.9	391	2 E72539	hypothetical prote
2	41	67.2	286	2 G90462	ABC transporter, p
3	40	65.6	662	2 T46005	receptor-like prot
4	39	63.9	105	2 H72708	probable ribosomal
5	39	63.9	436	2 G71862	alpha-(1,3)-fucosyl
6	39	63.9	454	2 B71914	alpha-(1,3)-fucosyl
7	39	63.9	476	2 C64601	fucosyltransferase
8	38	62.3	72	1 RSBPXL	excisionase - phag
9	38	62.3	72	2 S06533	excisionase - phag
10	38	62.3	72	2 A90729	excisionase [impor
11	38	62.3	201	2 T31302	hypothetical prote
12	38	62.3	230	2 H83001	probable permease
13	38	62.3	376	1 OXSE	beta-lactamase (EC
14	38	62.3	445	1 G70647	NMDH2 dehydrogenase
15	38	62.3	515	2 S20449	probable maturase,
16	38	62.3	515	2 S22550	hypothetical prote
17	38	62.3	717	2 T35219	probable membrane
18	37	60.7	141	2 H85217	hypothetical prote
19	37	60.7	179	2 T05810	hypothetical prote
20	37	60.7	182	2 T27243	hypothetical prote
21	37	60.7	220	2 D82726	conserved hypochet
22	37	60.7	732	2 S35651	cyclic nucleotide-
23	37	60.7	1355	2 T00075	hypothetical prote
24	36	59.0	98	1 G71488	conserved hypochet
25	36	59.0	223	2 T23356	hypothetical prote
26	36	59.0	275	2 S40007	trypsin (EC 3.4.21
27	36	59.0	411	2 B97340	competence-damage
28	36	59.0	506	2 F70617	probable aldehyde
29	36	59.0	524	1 O4RTWC	unspecific monooxy

30	36	59.0	561	2 G95367	probable methyl-ac
31	36	59.0	662	2 T44034	hypothetical prote
32	36	59.0	662	2 T44219	hypothetical prote
33	36	59.0	1189	2 T42726	guanine nucleotide
34	36	59.0	1226	2 H64479	protoporphyrin IX
35	36	59.0	1244	2 S29083	guanine-nucleotide
36	36	59.0	1260	2 S28407	guanine nucleotide
37	36	59.0	1275	2 A38985	nucleotide exchange
38	36	59.0	2115	2 S38480	nonstructural prot
39	36	59.0	2205	1 MNWVRN	nonstructural poly
40	36	59.0	2351	2 G71415	hypothetical prote
41	35	57.4	98	2 A75393	conserved hypochet
42	35	57.4	130	2 C87136	thioredoxin [impor
43	35	57.4	170	2 J70588	hypothetical 20K p
44	35	57.4	188	2 B72450	hypothetical prote
45	35	57.4	218	2 D90434	hypothetical prote

ALIGNMENTS

RESULT 1

E72539
hypothetical protein APE1602 - Aeropyrum pernix (strain K1)

C/Species: Aeropyrum pernix

C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #ext_change 20-Aug-1999

C/Accession: E72539

R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jhn-no, K.; Takai

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;

DNA, Res. 6, 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A/Reference number: A72450; WMD:99310339; PMID:10382966

A/Accession: E72539

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1391 <KAM>

A/Cross-references: DDBJ:AF000062; NID:G5105244; PIDN:BAA80602.1; PID:d1044388; PID:g51

A/Experimental source: strain K1

C/Genetics:

A/Gene: APE1602

Query Match 68.9%; Score 42; DB 2; Length 391;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVRRVVRWVR 12
Db 108 RVLRALERWVRK 119

RESULT 2

G90462
ABC transporter, permease (glucose) SSO2848 [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #ext_change 24-May-2001

C/Accession: G90462

R/She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.D.; Chan

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to Genbank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: G90462

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-286 <KUR>

A/Cross-references: GB:AE006641; NID:G13816206; PIDN:AAK42958.1; GSPDB:GN00155

C/Genetics:

A/Gene: SSO2848

Query Match 67.2%; Score 41; DB 2; Length 286;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;


```
.
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: APPLE MACINTOSH
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,171
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/225,476A
FILING DATE: 04-08-94
APPLICATION NUMBER: 08/039,620
FILING DATE: 06-04-93
APPLICATION NUMBER: 08/148,491
FILING DATE: 11-08-93
APPLICATION NUMBER: 08/148,889
FILING DATE: 11-08-93
ATTORNEY/AGENT INFORMATION:
NAME: HULTQUIST, STEVEN J.
REGISTRATION NUMBER: 28021
REFERENCE/DOCKET NUMBER: 4013-106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 37
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: NO
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-457-171-23

Query Match      27.8%; Score 60; DB 1; Length 37;
Best Local Similarity 26.5%; Pred. No. 0.08;
Matches          9; Conservative 15; Mismatches 10; Indels 0; Gaps 0;

Dn              7 RRVRRVRRVVRRVRVRRVRRVRRVRRVRRVRR 40
                :::::|::::|::::|::::|::::|::::|::::|
Db              3 KKFVKYAKVAKKAVAKKAVAKKAVAKKAVAK 36

RESULT 13
US-08-505-486-23
Sequence 23, Application US/08505486
Patent No. 5955573
GENERAL INFORMATION:
APPLICANT: Jesse M. Jaynes
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
NUMBER OF INVENTIONS: METHODS OF MAKING AND USING SAME
CORRESPONDENCE ADDRESS: 98
ADDRESS: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,486
FILING DATE: 21-JUL-1995
CLASSIFICATION: 536
```

[illegible]

CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-160

Query Match 28.7%; Score 62; DB 1; Length 28;
Best Local Similarity 42.9%; Pred. No. 0.034;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 12 RVRVRVRRVRRVRRVRRVRRV 39
DB 1 RVRVRVRRVRRVRRVRRVRRV 28

RESULT 8
US-08-932-682-160
Sequence 160, Application US/08932682
Patent No. 5945507
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mieczner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,682
FILING DATE: 18-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/786,748
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300

REFERENCE/DOCKET NUMBER: AP30421-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-705-5000
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5945507e
US-08-932-682-160

Query Match 28.7%; Score 62; DB 2; Length 28;
Best Local Similarity 42.9%; Pred. No. 0.034;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 12 RVRVRVRRVRRVRRVRRVRRV 39
DB 1 RVRVRVRRVRRVRRVRRVRRV 28

RESULT 9
US-08-231-730A-23
Sequence 23, Application US/08231730A
Patent No. 5561107
GENERAL INFORMATION:
APPLICANT: JAYNES, JESSE M.
APPLICANT: JULIAN, GORDON R.
TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST AN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: STEVEN J. HULTQUIST
ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
STREET: 200 PARK DRIVE, SUITE 210
STREET: P.O. BOX 14329
CITY: RESEARCH TRIANGLE PARK
STATE: NORTH CAROLINA
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: APPLE MACINTOSH
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231,730A
FILING DATE: 04-20-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-08-94
APPLICATION NUMBER: 08/039,620
FILING DATE: 06-04-93
APPLICATION NUMBER: 08/148,491
FILING DATE: 11-08-93
APPLICATION NUMBER: 08/148,889
FILING DATE: 11-08-93
ATTORNEY/AGENT INFORMATION:
NAME: HULTQUIST, STEVEN J.
REGISTRATION NUMBER: 28021
REFERENCE/DOCKET NUMBER: 4013-106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 37
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: NO
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/079,512
FILING DATE: 06/18/93
ATTORNEY/AGENT INFORMATION:
NAME: Roch, Michael J.
REGISTRATION NUMBER: 29,342
REFERENCE/DOCKET NUMBER: 0233 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 245-3594
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-179-632-9

Query Match 29.6%; Score 64; DB 1; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.021;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 18 RRVVRVRRVRRVRRVRRVRRV 45
DB 1 RRIYRAIRHPRIRGWLRIIGRIERV 28

RESULT 5
US-08-440-174A-9
Sequence 9, Application US/08440174A
Patent No. 5717061
GENERAL INFORMATION:
APPLICANT: Rao, Gururaj A.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: 7100 N.W. 62nd Avenue
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,174A
FILING DATE: 12-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/079,512
FILING DATE: 18-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bodrowicz, Donna
REGISTRATION NUMBER: 32,196
REFERENCE/DOCKET NUMBER: 0234R2D-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4896
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-174A-9

Query Match 29.6%; Score 64; DB 1; Length 31;

Best Local Similarity 42.9%; Pred. No. 0.021;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 18 RRVVRVRRVRRVRRVRRVRRV 45
DB 1 RRIYRAIRHPRIRGWLRIIGRIERV 28

RESULT 6
PCT-US95-00062-9
Sequence 9, Application PC/TUS9500062
GENERAL INFORMATION:
APPLICANT: Pioneer Hi-Bred International, Inc.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/Microsoft Windows
SOFTWARE: Microsoft Windows Notepad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00062
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Yates, Michael E.; Sweeney, Patricia A.;
NAME: Roch, Michael J.; & Simon, Soma G.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 234R2-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00062-9

Query Match 29.6%; Score 64; DB 5; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.021;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 18 RRVVRVRRVRRVRRVRRVRRV 45
DB 1 RRIYRAIRHPRIRGWLRIIGRIERV 28

RESULT 7
US-08-786-748A-160
Sequence 160, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mietzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:50:31, Search time 16.3404 Seconds
(without alignments)
86.430 Million cell updates/sec

Title: US-10-079-075-8
Perfect score: 216
Sequence: 1 RVRVRVRRVRRVRRVRRV.....RVRVRVRRVRRVRRVRR 48

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	33.8	38	2	US-08-436-703B-17
2	72	33.3	882	4	US-09-413-814-78
3	70	32.4	39	2	US-08-436-703B-5
4	64	29.6	31	1	US-08-179-632-9
5	64	29.6	31	1	US-08-440-174A-9
6	64	29.6	31	5	PCT-US95-00062-9
7	62	28.7	28	1	US-08-786-748A-160
8	62	28.7	28	2	US-08-932-682-160
9	60	27.8	37	1	US-08-231-730A-23
10	60	27.8	37	1	US-08-427-001C-23
11	60	27.8	37	1	US-08-457-798-23
12	60	27.8	37	1	US-08-457-171-23
13	60	27.8	37	2	US-08-505-486-23
14	60	27.8	37	3	US-08-689-489C-23
15	60	27.8	37	3	US-08-801-028-23
16	60	27.8	37	3	US-09-340-154-23
17	60	27.8	37	4	US-09-232-802A-23
18	60	27.8	37	4	US-09-482-611B-23
19	60	27.8	37	4	PCT-US94-061176-23
20	60	27.8	37	5	PCT-US94-12550-23
21	60	27.8	37	5	PCT-US95-04335-23
22	60	27.8	37	5	PCT-US95-04718-23
23	60	27.8	37	5	PCT-US95-09338-23
24	60	27.8	37	5	PCT-US95-09339-23
25	60	27.8	105	2	US-08-668-255-5
26	58	26.9	28	1	US-08-786-748A-18
27	58	26.9	28	2	US-08-932-682-18

28	58	26.9	96	2	US-08-668-255-7	Sequence 7, Appl
29	57	26.4	27	2	US-08-505-486-65	Sequence 65, Appl
30	57	26.4	27	3	US-08-801-028-65	Sequence 65, Appl
31	57	26.4	27	4	US-09-340-154-65	Sequence 65, Appl
32	57	26.4	27	3	US-09-482-611B-65	Sequence 65, Appl
33	57	26.4	27	5	PCT-US95-09338-65	Sequence 65, Appl
34	57	26.4	27	5	PCT-US95-09339-65	Sequence 65, Appl
35	57	26.4	28	1	US-08-786-748A-14	Sequence 14, Appl
36	57	26.4	28	1	US-08-786-748A-19	Sequence 14, Appl
37	57	26.4	28	1	US-08-786-748A-24	Sequence 14, Appl
38	57	26.4	28	1	US-08-786-748A-27	Sequence 14, Appl
39	57	26.4	28	1	US-08-786-748A-43	Sequence 14, Appl
40	57	26.4	28	1	US-08-786-748A-46	Sequence 14, Appl
41	57	26.4	28	1	US-08-786-748A-48	Sequence 14, Appl
42	57	26.4	28	2	US-08-932-682-14	Sequence 14, Appl
43	57	26.4	28	2	US-08-932-682-19	Sequence 14, Appl
44	57	26.4	28	2	US-08-932-682-24	Sequence 14, Appl
45	57	26.4	28	2	US-08-932-682-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-08-436-703B-17
Sequence 17, Application US/08436703B
Patent No. 5919761
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR
TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
TITLE OF INVENTION: WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benita J, Rohm, Esq.
STREET: 6601 Woodward Avenue
STREET: Suite 1525
City: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44mb, 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6;
SOFTWARE: ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7MK-060548-00233
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:57:02 ; Search time 62.1277 Seconds
(without alignments)
173.908 Million cell updates/sec

Title: US-10-079-075-8

Perfect score: 216
Sequence: 1 RVRVRVRRVRRVRRVRRV.....RRVRRVRRVRRVRRVRRV 48

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 1046584 seqs, 225093350 residues

Total number of hits satisfying chosen parameters: 1046584

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA.New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	35.2	220	US-10-425-114-68747	Sequence 68747, A
2	76	35.2	226	US-10-425-114-68746	Sequence 68746, A
3	76	35.2	230	US-10-425-114-68745	Sequence 68745, A
4	74	34.3	229	US-10-425-114-59258	Sequence 59258, A
5	69	31.9	28	US-09-874-644A-17	Sequence 17, Appl
6	65.5	30.3	181	US-10-424-599-20845	Sequence 20845, A
7	65.5	30.3	1374	US-10-366-683-24636	Sequence 24636, A
8	65.5	30.3	1374	US-10-419-128-24636	Sequence 24636, A
9	64.5	29.9	258	US-10-366-683-30124	Sequence 30124, A
10	64.5	29.9	128	US-10-419-128-30124	Sequence 30124, A
11	64	29.6	128	US-10-276-781-1853	Sequence 1853, Ap
12	62	28.7	205	US-10-156-761-14465	Sequence 14465, A
13	62	28.7	450	US-10-156-761-14465	Sequence 14465, A
14	61	28.2	79	US-10-017-161-2050	Sequence 2050, Ap
15	60.5	28.0	133	US-10-417-884-7073	Sequence 7073, Ap
16	60.5	28.0	133	PCT-US02-32727-13263	Sequence 13263, A
17	60.5	28.0	123	US-09-978-825-13263	Sequence 13263, A
18	60	27.8	321	US-10-057-498-13263	Sequence 13263, A
19	59.5	27.5	518	US-10-438-246-17097	Sequence 17097, A
20	59.5	27.5	518	US-10-366-683-21275	Sequence 21275, A
21	59.5	27.5	518	US-10-419-128-21275	Sequence 21275, A
22	59	27.3	874	US-10-438-246-18748	Sequence 18748, A
23	59	27.3	171	US-10-425-114-38005	Sequence 38005, A
24	58.5	27.1	261	US-10-425-114-71665	Sequence 71665, A
25	58.5	27.1	371	US-10-425-114-65011	Sequence 65011, A
26	58.5	27.1	6	US-10-446-203-10693	Sequence 10693, A
			648	US-10-366-683-24104	Sequence 24104, A

27	58.5	27.1	648	US-10-419-128-24104	Sequence 24104, A
28	58	26.9	179	US-09-675-784A-7224	Sequence 7224, Ap
29	58	26.9	331	US-10-282-122A-50182	Sequence 50182, A
30	58	26.9	742	US-10-366-683-29239	Sequence 29239, A
31	58	26.9	742	US-10-419-128-29239	Sequence 29239, A
32	58	26.9	1043	US-10-282-122A-50102	Sequence 50102, A
33	57.5	26.6	165	US-10-425-114-56931	Sequence 56931, A
34	57.5	26.6	612	US-10-424-599-161227	Sequence 161227, A
35	57.5	26.6	630	US-10-424-599-161230	Sequence 161230, A
36	57	26.4	365	US-10-282-122A-48107	Sequence 48107, A
37	57	26.4	381	US-10-425-114-65501	Sequence 65501, A
38	57	26.4	646	US-10-425-114-70188	Sequence 70188, A
39	57	26.4	720	US-10-282-122A-47999	Sequence 47999, A
40	57	26.4	759	US-60-426-500-4	Sequence 4, Appl
41	56.5	26.2	133	US-10-366-683-23880	Sequence 23880, A
42	56.5	26.2	133	US-10-419-128-23880	Sequence 23880, A
43	56.5	26.2	147	US-10-425-114-69755	Sequence 69755, A
44	56.5	26.2	151	US-10-425-114-38341	Sequence 38341, A
45	56.5	26.2	350	US-10-425-114-73059	Sequence 73059, A

ALIGNMENTS

```
RESULT 1
US-10-425-114-68747
; Sequence 68747, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68747
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Oryza sativa japonbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3598-036-C6_FLI.pep
US-10-425-114-68747

Query Match          35.2%; Score 76; DB 6; Length 220;
Best Local Similarity 33.8%; Pred. No. 1.4;
Matches 25; Conservative 11; Mismatches 12; Indels 26; Gaps 2;

QY 1 RVRVRVRRVRRVRRVRRV-----VRRVRRV-----RV 34
   |::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 73 RRRRLPRGLRLRLPRVLRRLGLPRLLAGRLRLPRVLRPRGLRRGV 132
   |::|::|::|::|::|::|::|::|::|::|::|::|
QY 35 VRRVRVRRVRRV 48
   |::|::|::|::|::|::|::|::|::|::|::|::|
DB 133 LRRACLPRVRR 146
   |::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 2
US-10-425-114-68746
; Sequence 68746, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68746
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Oryza sativa japonbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3598-036-C6_FLI.pep
US-10-425-114-68746
```

```
; Sequence 12, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
PCT-US02-04812-12
```

```
Query Match      78.7%; Score 170; DB 1; Length 48;
Best Local Similarity 90.9%; Pred. No. 1.1e-13;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 RVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 44
Db 1 RVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 44
```

```
RESULT 13
US-09-785-058-12
; Sequence 12, Application US/09785058
; GENERAL INFORMATION:
```

```
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-058-12
```

```
Query Match      78.7%; Score 170; DB 21; Length 48;
Best Local Similarity 90.9%; Pred. No. 1.1e-13;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 RVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 44
Db 1 RVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 44
```

```
RESULT 14
US-09-785-059-12
; Sequence 12, Application US/09785059
; GENERAL INFORMATION:
```

```
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-059-12
```

```
Query Match      78.7%; Score 170; DB 21; Length 48;
Best Local Similarity 90.9%; Pred. No. 1.1e-13;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 RVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 44
Db 1 RVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 44
```

```
RESULT 15
```

```
US-10-079-075-12
; Sequence 12, Application US/10079075
; GENERAL INFORMATION:
```

```
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-10-079-075-12
```

```
Query Match      78.7%; Score 170; DB 24; Length 48;
Best Local Similarity 90.9%; Pred. No. 1.1e-13;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 RVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 44
Db 1 RVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVRVVR 44
```

```
Search completed: June 9, 2003, 12:25:28
Job time : 207.83 secs
```

PCT-US02-04812-7
 Sequence 7, Application PC/US0204812
 GENERAL INFORMATION:
 APPLICANT: Ronald C. Montelaro
 APPLICANT: Timothy A. Miesner
 TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
 FILE REFERENCE: A34001-PCT / 072356.0223
 CURRENT APPLICATION NUMBER: PCT/US02/04812
 CURRENT FILING DATE: 2002-02-19
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 7
 LENGTH: 42
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURES:
 OTHER INFORMATION: Artificial peptide derived from HIV-1
 PCT-US02-04812-7

	Query Match	88.0%; Score 190; DB 1; Length 42;
	Best Local Similarity	100.0%; Pred. No. 3, 4e-16;
	Matches	42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	7 RRVRRVRRVVRRVVRRVVRRVVRRVVRRVVRR	48
Db	1 RRVRRVRRVVRRVVRRVVRRVVRRVVRRVVRR	42

```

RESULT 8
US-09-785-058-7
; Sequence 7, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Timothy A. Montelaro
; APPLICANT: Ronald C. Mieczner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
; US-09-785-058-7

```

	Query Match	88.0%	Score 190;	DB 21;	Length 42;
	Best Local Similarity	100.0%	Pred. No.	3,4e-16;	
	Matches 42;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	7 RRVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR	48			
	1 RRVVRRVVRRVVRRVVRRVVRRVVRRVVRRVVRR	42			

```

RESULT 9
US-09-785-059-7
/ Sequence 7, Application US/09785059
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A33577 / 072396.0217
/ CURRENT APPLICATION NUMBER: US/09/785,059
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 7
/ LENGTH: 42
/ TYPE: PRT
/ ORGANISM: Artificial sequence

```

```

; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-7

```

Qy	7	RRVRRVRRVRRVRRVRRVRRVRRVRRVRR	48
Db	1	RRVRRVRRVRRVRRVRRVRRVRRVRRVRR	42

```

RESULT 10
US-10-079-075-7
/ Sequence 7, Application US/10079075
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Mielezaro
/ APPLICANT: Timothy A. Mielezaro
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A34001-A / 072396.022
/ CURRENT APPLICATION NUMBER: US/10/079,075
/ CURRENT FILING DATE: 2002-02-19
/ NUMBER OF SEQ. ID NOS.: 12
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 7
/ LENGTH: 42
/ TYPE: PRT
/ ORGANSIM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-7

```

```

Oy      7 RRVRRRVRRRVRRRVRRRVRRRVRRRVRR 48
         |||||
Db      1 RRVRRRVRRRVRRRVRRRVRRRVRRRVRR 42
         |||||

Query Match      88.0%; Score 190; DB 24; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.4e-16;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 11
PCT-US02-04432-12
Sequence 12, Application PC/US0204432
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-PCT / 072396.0223
CURRENT APPLICATION NUMBER: PCT/US02/04432
CURRENT FILING DATE: 2002-02-13
NUMBER OF SEQ. ID NOS: 12
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 12
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: artificial peptides derived from HIV-1
PCT-US02-04432-12

```

	Query Match	78.7%;	Score 170;	DB 1;	Length 48;
	Best Local Similarity	90.3%;	Pred. No. 1.	Ie-13;	
	Matches	40;	Conservative	0;	Mismatches 4; Indels 0; Gaps 0;
Qy	1 RVRVVVRRVVRRVVRRVVRRVVRRVVRRVVRR	44			
Dz	1 RVRVVRRRVRRVVRRVVRRVVRRVVRRVVRRVVRR	44			

RESULT 12
PCT-US02-04812-12


```

; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-6

Query Match          75.0%; Score 162; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 8.1e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 48
DB 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 13
US-09-785-059-11
; Sequence 11, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-11

Query Match          62.0%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 2e-10;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 13 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 48
DB 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 14
US-10-079-075-11
; Sequence 11, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```

; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-11

Query Match          62.0%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 2e-10;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 13 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 48
DB 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 15
US-09-785-058-11
; Sequence 11, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-11

Query Match          62.0%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 2e-10;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 13 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 48
DB 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

Search completed: June 9, 2003, 12:34:10
Job time : 26.5319 secs
```



```

; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES

```


Query Match	28.7%	Score 62;	DB 20;	Length 28;
Best Local Similarity	42.9%	Pred. No. 0.24;		
Matches 12;	Conservative 8;	Mismatches 8;	Indels 0;	Gaps 0;

Qy	12	RVRRVRRVRRVRRVRRVRRVV	39
		: : : : :	
Db	1	RVRIRVGACRAIRHIIPRIRGRLRIL	28

RESULT 11
AAY56162
-ID AAY56162 standard; Protein; 172 AA

Query Match	28.7%;	Score 62;	DB 21;	Length 172;
Best Local Similarity	21.4%;	Pred. No. 1.4;		
Matches	12;	Conservative 17;	Mismatches 19;	Indels 8;
				Gaps 1;

Dy 1 RVRRVVRVVRRVRRVRR-----VRRVVRVVRRVVRVVRRVVRVVRR 48
||::||::||::||::|:::||:
Db 22 RVLHILEFVRKEEETMRKQGRMFEEVLRTLELMKKMEVEFKRVLHILEFVRK 77

RESULT	12
AAUS2068	
ID	AAUS2068 standard; Protein; 123 AA
XX	
AC	AAUS2068;

XX 27-FEB-2002 (first entry)
DT
XX
DE Propionibacterium acnes immunogenic protein #12964.

Query Match	28.0%	Score 60.5	DB 22	Length 123	.
Best Local Similarity	34.3%	Pred. No. 1.5			
Matches 23; Conservative	9	Mismatches 16	Indels 19	Gaps 3	

```
QY      1 RVVRRVVRRVVRKRV-----RVVRRVVVRVVRRVVR---RVVRRVVVRVV---- 42
        | :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB     21 RFRRLHRRITGHTIARQVPRVGVDRDRAHRLLTANARRVPRQGRGRHNRLVLRPRRQVGLL 80
```

QY	43	-RRVRR	48
		::	
Db	81	RRRIIR	87

RESULT 13

PT diseases and proliferative disorders such as cancer -
XX
PS Claim 1, Page 356-358; 392pp; English.
XX
CC The present invention is concerned with the sequence of the genome of the
CC SEN virus (SENV), and the proteins encoded by it. SENV is thought to be
CC the cause of hepatopathies which are not linked to the presence of the
CC hepatitis A, B and E viruses in man. The genome and proteins of this
CC virus can be used in gene therapy and vaccination against the virus,
CC which also causes disorders of the gastrointestinal tract, including
CC Crohn's disease and lupus erythematosus, inflammatory diseases, and
CC proliferative disorders such as cancer.
XX
SQ Sequence 743 AA;
XX
Query Match 29.2%; Score 63; DB 21; Length 743;
Best Local Similarity 53.3%; Pred. No. 4.4;
Matches 24; Conservative 2; Mismatches 11; Indels 8; Gaps 3;
XX
QY 4 RVRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVR 48
34 RPYRR--RQRRVRVRFRGRGRGR--RRYIR-----RRRLR 70
XX
RESULT 9
AAW47769
ID AAW47769 standard; peptide; 28 AA.
XX
XX AAW47769;
XX
DT 26-MAY-1998 (first entry)
XX
DE Antimicrobial peptide LLPI analogue.
XX
XX Antimicrobial; transmembrane protein; TM; lentivirus lytic peptide;
XX LLPI; amphipathic; antibacterial; antifungal; antiviral; antiprotoczoal.
XX
OS Synthetic.
OS Human immunodeficiency virus.
XX
PN US5714577-A.
XX
PD 03-FEB-1998.
XX
PF 24-JAN-1997; 97US-0786748.
XX
PR 26-JAN-1996; 96US-0010634.
PR 24-JAN-1997; 97US-0786748.
XX
XX (UYPI-) UNIV PITTSBURGH.
XX
PI Metzner TA, Montelaro RC, Tencza SB;
XX WPI; 1998-158352/14.
XX
PT Retroviral TM peptides - useful as antibacterial agents
XX
PS Disclosure; Column 19; 59pp; English.
XX
CC The invention relates to new antimicrobial peptides which correspond to
CC amino acid sequences in the transmembrane proteins of lentiviruses, in
CC particular HIV and SIV. These peptides comprise arginine rich sequences
CC which, when modelled for secondary structure, display high
CC amphipathicity and hydrophobic moment. Also disclosed are structural
CC and functional analogues and homologues of these peptides which also
CC display antimicrobial activity. The peptides are highly inhibitory to
CC microorganisms (bacteria, fungi, viruses and protozoa) but significantly
CC less toxic to red blood cells and other normal mammalian cells. Activity
CC is demonstrated against Gram positive and negative bacteria including
CC *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Enterococcus faecalis* and
CC *Serratia marcescens*.
CC The present sequence is one of 169 disclosed specific examples of
CC the new peptides. It is an analogue of the peptide designated LLPI

CC (see AAW47614) which is a peptide from the transmembrane protein (gp41)
CC of HIV strain HXB2K.
XX
SQ Sequence 28 AA;
XX
Query Match 28.7%; Score 62; DB 19; Length 28;
Best Local Similarity 42.9%; Pred. No. 0.24;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
XX
QY 12 RVRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVR 39
1 RVRVRQACRAIRHPRIRQSLRRL 28
XX
RESULT 10
AAV32703
ID AAV32703 standard; peptide; 28 AA.
XX
XX AAV32703;
XX
DT 21-OCT-1999 (first entry)
XX
DE Antimicrobial peptide LLPI analogue.
XX
XX Antimicrobial peptide; LLPI; SLP-1; LIP2; SLP2A; SLP2B; ELP; infection;
XX growth inhibitor; microorganism; virus; gene therapy; vector production;
XX sterilisation.
XX
OS Synthetic.
OS Human immunodeficiency virus type 1.
XX
PN US5945507-A.
XX
PD 31-AUG-1999.
XX
PF 18-SEP-1997; 97US-0932682.
XX
PR 26-JAN-1996; 96US-0010634.
PR 24-JAN-1997; 97US-0786748.
PR 18-SEP-1997; 97US-0932682.
XX
XX (UYPI-) UNIV PITTSBURGH.
XX
PI Metzner TA, Montelaro RC, Tencza SB;
XX WPI; 1999-508189/42.
XX
PT Antimicrobial peptides useful for treating microbial infections
XX
PS Disclosure; Column 21; 62pp; English.
XX
CC This sequence represents an antimicrobial peptide of the invention, and
CC is an analogue of the peptide LLPI (see AAV32549). The peptides can be
CC used for treating infections caused by *Staphylococcus aureus*,
CC methicillin resistant *S. aureus*, *Pseudomonas aeruginosa*, *Enterococcus*
CC *faecalis*, *S. marcescens*, *Escherichia coli*, fungi, protozoa and viruses in
CC a mammalian host. They can be used to inhibit growth of diverse
CC microorganisms such as bacteria, fungi, protozoa and DNA and RNA viruses
CC and can be used in tissue culture to inhibit unwanted microbial growth,
CC particularly for the production of recombinant proteins or vectors for
CC gene therapy. They can also be used in preventing infections through the
CC sterilisation of wounds prior to suture and to sterilise surgical
CC instruments. The unique structure of these antimicrobial peptides
CC imparts high potency while selectivity is maintained, they are
CC moderately haemolytic but only lyse red blood cells at high
CC concentrations unlike melittin, a peptide extracted from bee venom, which
CC is highly active against bacteria and lyses red blood cells showing
CC little selectivity. The peptides target a membrane structure which makes
CC it more difficult for a microorganism to develop a mechanism of
CC resistance against this type of antibiotic. Their small size makes them
CC relatively simple to prepare by standard synthetic peptide chemistry.
XX
SQ Sequence 28 AA;

KM	Antimicrobial; antifungal; pathogen; plant; amphipathic;
KW	broad spectrum.
XX	
OS	Synthetic.
XX	
PN	W09S18855-A2.
XX	
PD	13-JUL-1995.
XX	
PP	06-JAN-1995; 95WO-US00062.
XX	
PR	07-JAN-1994; 94US-0179632.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
P1	Rao AG, Zhong L;
XX	
DR	WPI, 1995-255059/33.
XX	
PT	New antimicrobial amphipathic polypeptide(s) and related nucleic
PT	acids - for clinical use or esp. to increase resistance of plants
XX	to fungal pathogens.
XX	
PS	Claim 1; Page 17; 24pp; English.
XX	
CC	AA80727-R80748 are new amphipathic polypeptides. They have a broad
CC	spectrum of antimicrobial and antifungal activity. They can be used
CC	to treat or prevent infection in humans and animals or applied to
CC	plants as sprays, creams, dust, etc. The DNA encoding these
CC	peptides can also be incorporated into susceptible plants via the
CC	use of a non-phytotoxic vehicle adapted for systemic administration.
CC	This process imparts resistance to plant pathogens esp. fungi (e.g.
CC	Fusarium graminearum, F.molinitforme, Aspergillus flavus, Alternaria
CC	longipes, Colletotrichum graminicola, Phytophthora megasperme,
CC	Sclerotinia sclerotium). The peptides are esp. useful in
CC	transformed plants such as maize, sorghum, wheat, soya, alfalfa,
CC	rapeseed, sunflower, tobacco or tomato.
XX	
SQ	Sequence 31 AA:
Query Match	29.6%; Score 64; DB 16; Length 31;
Best Local Similarity	42.9%; Pred. NO. 0.15;
Matches 12; Conservative	7; Mismatches 9; Indels 0; Gaps 0;
OY	18 RRVVVVVVVVVVVVVVVVVVVVVVVVV 45
DB	1 RRIYRAIRHPIRRIRIGWLRIGRIRERY 28
RESULT 7	
AA024328	
AA024328	standard; Protein; 128 AA.
12-OCT-2001	(first entry)
C parvum	EST encoded protein SEQ ID NO: 1853.
Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;	
tomato; monkey; dog; sea urchin; expressed sequence tag; EST;	
diagnostics; forensic test; gene mapping; genetic disorder;	
biodiversity; gene therapy; nutrition.	
Cryptosporidium parvum.	
25-JAN-2001; 2001WO-US02687.	
25-JAN-2000; 2000US-0491404.	

	XX	PA	(HYSE-) HYSEQ INC.
	XX	PB	Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
	PI	PCao Y,	Dremanac RA, Zhang J, Werhan T;
	XX	PD	WPI; 2001-476164/51.
	DR	N-PSDB;	AHH98987.
	XX	PT	Isolated polypeptide for treatment of diseases, diagnostics, raising antibodies and research use -
	XX	PS	Claim 20; Page 1203; 1275pp; English.
	CC	The present invention provides the protein and coding sequences of novel proteins from a variety of organisms, including human, dog, cat, horse, cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea urchin and tomato. These were derived from expressed sequence tags (ESTs) from the organism of interest. They can be used in diagnostics, forensics, gene mapping, identification of mutations, to assess biodiversity and for nutritional purposes. The present sequence is a protein of the invention.	
	CC	SQ	Sequence 128 AA;
	Query Match	Best Local Similarity	29.6%; Score 64; DB 22; Length 128; Matches 27; Conservatvie 3; Mismatches 9; Indels 8; Gaps 3;
OY	2	VVVVRRRVRRRVRVVVRRRVRRRVRR-----RVVRRV	43 ::: :: :: :: :: :: :: :: :: :: :
Db	3	VVVTV--VVVVVTAVV-RVMVVVVVVVVVVVVVVVVVVVVVVVVVVVV	46
RESULT 8			
ID	AAH1535	standard; Protein; 743 AA.	
AC	AAH1535;		
DT	19-DEC-2000	(first entry)	
DE	SEN virus protein fragment SEQ ID NO: 122.		
KM	SEN virus; SENV; gastrointestinal tract disorder; inflammatory disease; proliferative disorder; hepatopathy; hepatitis; viral infection; vaccination; gene therapy.		
OS	Hepatitis virus.		
PN	WO20028039-A2.		
PD	18-MAY-2000.		
PF	09-NOV-1999;	99WO-EPO8566.	
PR	10-NOV-1998;	98IT-MI02437.	
PR	30-APR-1999;	99IT-MIO0923.	
PR	14-MAY-1999;	99EP-O830298.	
PR	16-JUL-1999;	99EP-O113932.	
PA	(DIAS-) DIASORIN SRL.		
PI	Primi D, Fiordalisi G, Mantero GL, Mattioli S, Scattini A;		
PI	Bonelli F, Vaglini L, Olivero P, Dal Corso A, Bonelli M;		
DR	WPI; 2000-376551/32.		
FT	Nucleic acids representing the genome of the SEN virus (SENV) and encoded proteins, useful for treatment of hepatopathies, inflammation		

[illegible]

SQ	Sequence	18 AA;
Query Match	31.0%; Score 67; DB 22; Length 18;	
Best Local Similarity	47.1%; Pred. No. 0.039; Indels 0; Gaps 0;	
Matches	8; Conservative 9; Mismatches 0; Indels 0; Gaps 0;	
OY	15 RVRRVVRVVRRVVRRV 31	
	:: :: :: :: :: :	
DB	2 RIIRRIIRRIIRRI 18	
RESULT 5		
AAR60065		
ID	AAR60065 standard; peptide; 31 AA.	
XX		
AC	AAR60065;	
XX		
DT	16-FEB-1995 (first entry)	
XX		
DE	Antimicrobial peptide.	
XX		
KW	Amphipathic peptide; alpha-helix; lytic peptide; antifungal;	
KM	antimicrobial; fungus resistance; disease resistance;	
KW	crop protection.	
XX		
OS	Synthetic.	
XX		
PN	WO9415961-A.	
XX		
PD	21-JUL-1994.	
XX		
PF	12-JAN-1994; 94WO-US00383.	
XX		
PR	13-JAN-1993; 93US-0003864.	
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Rao AG, Zhong L;	
XX		
DR	WPI, 1994-249137/30.	
XX		
PT	Synthetic polypeptide(s) and the nucleic acid encoding them -	
PT	exhibits amphipathic alpha-helices and provide cell-expressable	
PT	antimicrobial activity	
XX		
PS	Disclosure; Page 19; 25pp; English.	
CC	The synthetic amphipathic alpha-helical lytic peptides given in	
CC	AAR60057-71 were designed to provide antifungal or antimicrobial	
CC	activity when expressed in monocot or dicot plants. They also have	
CC	veterinary and medical applications.	
XX		
SQ	Sequence 31 AA;	
Query Match	29.6%; Score 64; DB 15; Length 31;	
Best Local Similarity	42.9%; Pred. No. 0.15; Indels 0; Gaps 0;	
Matches	12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;	
OY	18 RRVVRVVRRVVRRVVRRV 45	
	:: :: :: :: :: :	
DB	1 RRIYRAIRHIPPRIIRIGMLRIGRIERI 28	
RESULT 6		
AAR80735		
ID	AAR80735 standard; peptide; 31 AA.	
XX		
AC	AAR80735;	
XX		
DT	28-FEB-1996 (first entry)	
XX		
DE	Synthetic antimicrobial/antifungal polypeptide.	

DR WPI; 1995-276981/37.
 XX Complex of nucleic acid and oligopeptide with sec. structure - and
 PT transfer vectors contg. them, useful for efficient transfer of
 PT nucleic acid to cells in gene therapy.
 XX
 PS Claim 6; Page 16; 20pp; French.
 XX
 CC The present peptide corresponds to a generic formula for a cationic
 CC oligopeptide; the formula is (b-1-1-b)n, where b is a hydrophobic amino
 CC acid, 1 is a hydrophilic amino acid and n is at least 4. In this case,
 CC where b is Leu, 1 is Arg and n = 10, the oligopeptide forms an alpha-
 CC helix which forms a stable complex with a nucleic acid. The complex
 CC is suitable for transferring nucleic acid, esp. in gene therapy.
 CC
 SQ Sequence 40 AA;
 Query Match 35.4%; Score 76.5; DB 16; Length 40;
 Best Local Similarity 46.2%; Pred. No. 0.0062;
 Matches 18; Conservative 18; Mismatches 0; Indels 3; Gaps 3;
 QY 13 VRRVRRVV-RVRRVRRV-RVRRVRRV-RVRRVRRV 48
 Db 1 LRRRLRLRLRLRLRLRLRLRLRLRLRLRLRLRLRL 39
 RESULT 2
 AAM06684
 ID AAM06684 standard; peptide; 39 AA.
 AC AAM06684;
 XX
 DT 05-AUG-1997 (first entry)
 XX
 DE Protamine-like peptide analogue [+18RGD].
 KW Coagulation; anticoagulant; heparin; platelet aggregation; cell adhesion;
 KM positively charged cluster; arginine; polycationic; decrease;
 XX n-protamine; salmine protamine; protamine sulphate; salmon sperm.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT Region 3..30
 FT Binding-site /note= repeat_region
 FT Modified-site /note= "4 tandem repeats of (Arg)2(Ala)2Ala motif"
 FT /label= cell_adhesion_motif
 FT /note= "amidated"
 XX
 XX W09635444-A1.
 XX 14-NOV-1996.
 PD
 XX 08-MAY-1996; 96MO-US06567.
 PF
 XX 08-MAY-1995; 95US-0436703.
 PR
 XX (UNMI) UNIV MICHIGAN.
 PA
 XX Andrews PC, Stanley JC, Wakefield TW;
 PI WPI; 1997-011697/01.
 DR
 XX Peptide reversing the anticoagulant effects of heparin - is based on
 PT protamine but has fewer positive charges for reduced toxicity
 XX
 PS Claim 31; Page 31; 42pp; English.
 CC Protamine sulphate (also called n-protamine or salmine protamine) is

CC a polycationic peptide derived from salmon sperm and is used to
 CC reverse heparin anticoagulation. One of the major components of
 CC salmine protamine is a 32 amino acid peptide having a total cationic
 CC charge of [+21], with arginine accounting for 67% of the total sequence
 CC and for all of the positive charge. Peptides of 20-40 amino acids with
 CC total cationic charge less than [+21] and which are able, at least
 CC partially, to reverse the effect of heparin and/or low molecular weight
 CC heparin anticoagulants are claimed. Specifically, the peptides are
 CC polycationic analogues of n-protamine where the positive charge on the
 CC amino acid sequence is reduced by selective replacement of positively
 CC charged arginine residues with an uncharged residue, so that total
 CC cationic charge is less than [+21]. The new peptides are used in vivo
 CC to reverse the effects of heparin; they have the same anti-heparin
 CC activity as protamine but are less toxic (because of the reduced
 CC number of positive charges) and are relatively easy and inexpensive
 CC to prepare. The present sequence represents a specifically claimed
 CC protamine-like peptide with a charge of [+18]; this peptide also
 CC includes an RGD fibronectin receptor ligand motif. Peptide [+18RGD]
 CC improved the reversal of anti-Factor Xa activity of the low molecular
 CC weight heparin Enoxaparin to 72%, compared to only 30% reversal by
 CC protamine. Also, peptide [+18RGD] produced less decrease in platelet
 CC count than n-protamine.
 CC
 SQ Sequence 39 AA;
 Query Match 32.4%; Score 70; DB 18; Length 39;
 Best Local Similarity 50.0%; Pred. No. 0.036;
 Matches 17; Conservative 0; Mismatches 13; Indels 4; Gaps 1;
 QY 7 RRVRRRRRVRRVRRVRRVRRVRRVRRVRRVRRV 40
 Db 3 RRAARRARRARRR---ARRARRARRARRARRR 32
 RESULT 3
 AAP91336
 ID AAP91336 standard; peptide; 28 AA.
 AC AAP91336;
 XX
 DT 19-MAR-1990 (first entry)
 XX
 DE Amino acid sequence of Shiva-4.
 KW Shiva-4; lytic peptide; antimicrobial peptide; disease-resistant
 KM trichophyte; Shiva-2; Shiva-3; Shiva-5; Shiva-6; Shiva-7.
 XX
 XX W08904371-A.
 XX
 PD 18-MAY-1989.
 PF 02-NOV-1988; 88MO-US03908.
 XX
 XX 02-NOV-1987; 87US-0115941.
 PR
 XX (LOU) LOUISIANA STATE UNIV.
 PA
 XX Jaynes JM, Derrick KS;
 PI WPI; 1989-165650/22.
 DR
 XX Transformed plants contg. heterologous gene - expressing antimicrobial
 PT agent, or polypeptide high in essential amino acids
 CC
 XX Table I; 56pp; English.
 PS
 CC Amino acid sequence of Shiva-4 as an exemplary lytic peptide for
 CC use as an antimicrobial peptide contemplated for use in plant
 CC (trichophyte) transformants in the invention. It is a homologue of
 CC Shiva-2, -3 and -5 to -7. All of these Shiva peptides are also
 CC contemplated as having general utility in inducing lysis of cells in
 CC vitro. Shiva-4 may be too lytically active to be used in plants at high
 CC expression levels.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:43:01 ; Search time 44.0851 Seconds

(without alignments)
224.345 Million cell updates/sec

Title: US-10-079-075-8

Perfect score: 216
Sequence: 1 RVRVVRVVRVVRVVRVVRV.....RVRVVRVVRVVRVVRV 48

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	37.0	171	12 071097	071097 bovine aden
2	76	35.2	9	094M13	094M13 bacterioph
3	74.5	34.5	349	10 095HX2	095HX2 arbidopsis
4	74	34.3	428	17 08TXS5	08TXS5 methanopyru
5	70	32.4	230	9 09FZV7	09FZV7 pseudomonas
6	68.5	31.7	302	5 025562	025562 naegleria g
7	67.5	31.2	769	12 099AR5	099AR5 tt virus. o
8	65.5	30.3	759	12 091CZ2	091CZ2 tt virus. o
9	65	30.1	114	12 084529	084529 paramecium
10	65	30.1	184	9 064325	064325 bacterioph
11	64	29.6	349	16 0987V7	0987V7 rhizobium 1
12	64	29.6	682	16 08XVM2	08XVM2 raietonia s
13	64	29.6	742	12 091PS3	091PS3 tt virus. o
14	62.5	28.9	115	12 09W9G8	09W9G8 human herpe
15	61.5	28.5	85	2 09Z1Y4	09Z1Y4 borrelia bu
16	61	28.2	74	16 09ZK12	09ZK12 helicobacte

17	61	28.2	211	12 069587	069587 human herpe
18	61	28.2	321	12 069545	069545 human herpe
19	61	28.2	311	12 069581	069581 human herpe
20	61	28.2	635	2 030348	030348 raietonia s
21	60.5	28.0	651	10 039720	039720 euglena gra
22	60	27.8	105	5 025305	025305 leishmania
23	60	27.8	105	5 097Y18	097Y18 leishmania
24	60	27.8	650	10 039721	039721 euglena gra
25	60	27.8	755	12 091D00	091D00 tt virus. o
26	60	27.8	862	15 090CG7	090CG7 human immun
27	59.5	27.5	331	17 08TXA5	08TXA5 methanopyru
28	59.5	27.5	513	16 09R205	09R205 streptomyce
29	59.5	27.5	761	12 09WT88	09WT88 tt virus. o
30	59	27.3	263	10 09LT00	09LT00 arbidopsis
31	58.5	27.1	100	5 09GRP6	09GRP6 leishmania
32	58.5	27.1	100	5 000912	000912 leishmania
33	58.5	27.1	100	5 000913	000913 leishmania
34	58.5	27.1	166	3 005385	005385 saccharomyc
35	58.5	27.1	217	5 09V6X3	09V6X3 drosophila
36	58.5	27.1	619	5 095S18	095S18 drosophila
37	58.5	27.1	723	12 09DUC4	09DUC4 tt virus. o
38	58	26.9	94	12 08V7D3	08V7D3 tt virus. o
39	58	26.9	138	16 09RD30	09RD30 streptomyce
40	58	26.9	312	5 021320	021320 caenorhabdi
41	58	26.9	389	17 08TX11	08TX11 methanopyru
42	58	26.9	738	12 08V7H2	08V7H2 tt virus. o
43	58	26.9	842	15 09QM83	09QM83 human immun
44	58	26.9	855	15 08UL67	08UL67 human immun
45	58	26.9	869	15 08UL69	08UL69 human immun

ALIGNMENTS

RESULT 1

071097 ID 071097 PRELIMINARY; PRT; 171 AA.

AC 071097;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE PY11.
 OS Bovine adenovirus type 3 (Mastadenovirus boes3).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10510;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WB-1;
 RX MEDLINE=98105785; PubMed=9445040;
 RA Reddy P.S., Idamakanti N., Zakharichouk A.N., Baxi M.K., Lee J.B.,
 RA Pyne C., Babiuk L.A., Tikoo S.K.;
 RT "Nucleotide sequence, genome organization, and transcription map of
 RT bovine adenovirus type 3";
 RL J. Virol. 72:1394-1402 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WB-1;
 RX MEDLINE=98318755; PubMed=9654686;
 RA Baxi M.K., Reddy P.S., Zakharichouk A.N., Idamakanti N., Pyne C.,
 RA Babiuk L.A., Tikoo S.K.;
 RT "Characterization of bovine adenovirus type 3 early region 2B";
 RL Virus Genes 16:313-316 (1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WB-1;
 RX MEDLINE=98451815; PubMed=9778793;
 RA Lee J.B., Baxi M.K., Idamakanti N., Reddy P.S., Zakharichouk A.N.,
 RA Pyne C., Babiuk L.A., Tikoo S.K.;
 RT "Genetic organization and DNA sequence of early region 4 of bovine
 RT adenovirus type 3";
 RL Virus Genes 17:99-100 (1998).
 RN [4]
 RP SEQUENCE FROM N.A.

RA Haelekom R., Kyriades N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
Brucella melitensis";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -1- FUNCTION: The ruva-ruvb complex in the presence of ATP renatures
CC chromatin structure in supercoiled DNA with palindromic sequence,
CC indicating that it may promote strand exchange reactions in
CC homologous recombination. RuvAB is an helicase that mediates the
CC Holliday junction migration by localized denaturation and
CC reannealing (By similarity).
CC -1- SUBUNIT: Forms a complex with ruva (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RUVB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE009475; AAL51515.1; -
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR003959; AAA_ATPase_cent.
CC InterPro: IPR004605; RUVB.
CC Pfam: PF00004; AAA; 1.
CC TIGRFAMs: TIGR00635; ruvb; 1.
CC DNA repair: SOS response; ATP-binding; DNA recombination; Helicase;
CC Complete proteome.
CC NP_BIND 59 ATP (POTENTIAL).
CC SEQUENCE 346 AA; 38268 MW; F97710688F919FBC CRC64;

Query Match : 25.7%; Score 55.5; DB 1; Length 346;
Best Local Similarity 34.6%; Pred. NO. 5.4;
Matches 18; Conservative 7; Mismatches 18; Indels 9; Gaps 1;

Qy 3 VVVVRRVRRVRRVRRVRRV-----RRVRRVRRVRRVRRVRRV 45
Db 176 VRLNPFYEELEYIVRGARLTMQNGISSDGAREVARRSGTPRIYGRLLRV 227

RESULT 14
L2MU_ADEB2
ID L2MU_ADEB2 STANDARD; PRT; 70 AA.
AC Q96626;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Late L2 mu core protein precursor (pmu) (Protein X).
GN PX.
OS Bovine adenovirus type 2 (Mastadenovirus b0s2).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=114429;
RN [1]
RP SEQUENCE FROM N.A.
RA Ruvaeli M., Harrach B., Banreyl A., Evans P., Benko M.;
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: THE ROLE OF THE PRECURSOR MIGHT BE TO CONDENSE THE
CC VIRAL PROCHROMATIN FOR ENCAPSIDATION BY VIRTUE OF THE TWO
CC BASIC DOMAINS (BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U44123; AAB16758.1; -
CC Core protein; DNA-binding; Late protein.
CC PROPEP 1 26 BY SIMILARITY.
FT CHAIN 27 41 LATE L2 MU CORE PROTEIN.

[illegible]

RA "Complete genome sequence of the alkaliophilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
RC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on items
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isdb.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL; AP001511; BAB04994.1; -.
DR InterPro; IPR001324; PRK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIDBKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KM Transferrase; Kinase; ATP-binding; Complete proteome.
FT NP_BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 211 AA; 24387 MW; C2AFACB0030520B CRC64;

Query Match 25.9%; Score 56; DB 1; Length 211;
Best Local Similarity 37.9%; Pred. No. 2.9;
Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

DY 15 RVRRRVRRVRRRVRRRVRRVVRR 43
DB 138 RIIRRYVDRIERGRTLESVEIQYKKVR 166
||::|::|::|::|::|:
PRTI CLUPA STANDARD; PRT; 30 AA.
ID ID_PRTI CLUPA
AC P02335;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Protamine YII (Clupeine YII).
OS Clupea pallasii (Pacific herring), and
OS Clupea harengus (Atlantic herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
OC Clupea.
OX NCBI_TaxID=30724, 7950;
RN RN [1]
RP SEQUENCE.
RC SPECIES=C.pallasii;
RX MEDLINE=73223106; Pubmed=4664740;
RA Suzuki K., Ando T.;
RT "Studies on protamines. XVI. The complete amino acid sequence of
RT clupeine YII.";
RJ J. Biochem. 72:1419-1432 (1972).
RN RN [2]
RP SEQUENCE.
RC SPECIES=C.harengus;
RA Chang W.J., Nukushina M., Ishii S., Nakahara C., Ando T.;
RA Submitted (Aug-1970) to the PIR data bank.
CC -1- FUNCTION: PROLAMINE SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC PIR; A02677; CMHRX2.
CC PIR; A37575; CMHRZ2.
KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;

```

QY      Query Match          25.7%; Score 55.5; DB 1; Length 30;
        Best Local Similarity 54.3%; Pred. No. 0.49;
        Matches 19; Conservative 0; Mismatches 9; Indels 7; Gaps 2.
DB      14 RRVRVRRVRRVRRVRRVRRVRRVRRVRRVRR 48
        |||||
        2 RRRTRRASRPVRR--RRRPRVSRR-----RRARR 29

RESULT 12
SSSL_SCYC ID SSSL_SCYC STANDARD; PRT; 87 AA.
AC PI3275;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 01-JAN-1990 (Rel. 13, Last annotation update)
OS Spermatid-specific protein Sl.
OC Scyllorhinus canicula (Spotted dogfish) (Spotted catchark);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scylliorhiniidae; Scylliorhinus;
OX NCB1_TaxID=7830;
RN [1]
RP SEQUENCE.
RX MEDLINE=88055030; PubMed=3678229;
RA Chauviere M., Martinage A., Briand G., Sauciere P., Chevallier P.;
RT "Nuclear basic protein transition during sperm differentiation. Amino
RT acid sequence of a spermatid-specific protein from the dog-fish
RT Scylliorhinus caniculus."
RL Eur. J. Biochem. 169:105-111(1987).
CC -1- FUNCTION: INVOLVED IN NUCLEAR BASIC PROTEIN TRANSITION: HISTONES
CC ARE REPLACED BY SPERMATID SPECIFIC PROTEINS WHICH ARE THEMSELVES
CC REPLACED BY PROTAMINES IN LATE SPERMATIDS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: WITH PROTEIN S2.
DR PIR; S00180; S00180.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Nuclear protein.
SQ SEQUENCE 87 AA; 11179 MW; 38A0ED7A82536910 CRC64;

Query Match          25.7%; Score 55.5; DB 1; Length 87;
Best Local Similarity 41.9%; Pred. No. 1.4;
Matches 18; Conservative 6; Mismatches 12; Indels 7; Gaps 2;
QY      7 RRVRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 44
        |||||
DB      14 RRRGRRRKKT--RCRRKGRRISRRPRTTYYRRRVKKIVHLKR 54

RESULT 13
RUVB_BRUME
RD_RUVB_BRUME STANDARD; PRT; 346 AA.
AC O8YT5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Holiday junction DNA helicase ruvb.
DE RUVB OR BMEI0334.
GN Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OC NCBI_TaxID=29459;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapetral V., Redkar R.J., Patra G., Mujer C., Jos T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lytkids A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Leifsson J.-O.,
```



```

RESULT 5
HSP2_HORSE
ID      HSP2_HORSE      STANDARD;      PRT;      62 AA.
AC      P15342;
DT      01-APR-1990 (Rel..14, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Sperm histone p2a (ST2A).
OS      Equus caballus (Horse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX      NCBI_TaxID=9796;
RN      [1]
RP      SEQUENCE.
RX      MEDLINE=90304188; PubMed=2364093;
RA      Pithonen A., Valonen P., Linnala-Kankkunen A., Heiskanen M.-L.,
RA      Maenpaa P.K.;
RT      "Primary structures of two protamine 2 variants (St2a and St2b) from
RT      stallion "spermatzoa.";
RL      Biochim. Biophys. Acta 1039:177-180(1990).
RN      [2]
RP      SEQUENCE OF 1-25.
RX      MEDLINE=89171259; PubMed=2924903;
RA      Pithonen A., Linnala-Kankkunen A., Maenpaa P.K.;
RT      "Comparison of partial amino acid sequences of two protamine 2
RT      variants from stallion sperm. Structural evidence that the variants
RT      are products of different genes.";
RL      FEBS Lett. 244:199-202(1989).
CC      -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC      SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC      SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC      -1- SUBCELLULAR LOCATION: Nucleus.
CC      -1- TISSUE SPECIFICITY: TESTIS.
CC      -1- SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.
DR      PIR; S10754; S10754.
DR      PIR; S10754; S10754.
KW      Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW      Testis; DNA condensation; Nuclear protein.
SQ      SEQUENCE      62 AA;      8418 MW;      15EABCI09F2AD6B1 CRC64;

Query Match:      27.3%;      Score 59;      DB 1;      Length 62;
Best Local Similarity      50.0%;      Pred. No. 0.4;
Matches      22;      Conservative      2;      Mismatches      18;      Indels      2;      Gaps      1;

QY      1 RVEVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 44
DB      20 RLYELRRRRVRRSSRRRRRRCOR--RRRRRRVRCFRRRRRRCRR 61

RESULT 6
VC07_ADEB2
ID      VC07_ADEB2      STANDARD;      PRT;      183 AA.
AC      Q96624;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Major core protein precursor (Protein VII) (pVII).
OS      Bovine adenovirus type 2 (Mastadenovirus boe2).
OC      Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX      NCBI_TaxID=11442;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Ruuvola M., Harrach B., Banreivi A., Evans P., Benko M.;
RA      Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial

```

[illegible]


```

RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RX MEDLINE=94087748; PubMed=8263936;
RA Davidson A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
RT "The DNA sequence of adenovirus type 40."
RN J. Mol. Biol. 234:1308-1316(1993).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RA Pleniakzek N.J., Slemenda S.B., Pleniakzek D., Luftig R.B.;
RL Submitted (FEB-1992) to the EMBL/Genbank/DBJ databases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ibt-sib.ch/announce/
CC or send an email to license@ibt-sib.ch).
CC -----
CC EMBL; L19443; AAC13963.1; -.
DR EMBL; M86665; AAA42526.1; -.
DR InterPro; IPRO04912; Adeno VII.
DR Pfam; PF03228; Adeno VII; 1.
DT PROPEP 1 23 BY SIMILARITY.
FT CHAIN 24 185 MAJOR CORE PROTEIN.
FT SITE 23 24 CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT SITE (POTENTIAL).
SQ SEQUENCE 185 AA; 20518 MW; 4FB0B535F218A9E CRC64;
-----
OY Query Match 29.6%; Score 64; DB 1; Length 185;
Db Best Local Similarity 37.8%; Pred. No. 0.32;
Matches 17; Conservative 7; Mismatches 15; Indels 6; Gaps 1;
2 VVRVVRVVRVVRVVRVVR-----VVRVVRVVRVVRVVRVVRVVR 40
87 VVANRRYQKRRLLQRRRRPPTAARAVLRRQRTGRRAVR 131
-----
RESULT 3
HSP2_ALOSE STANDARD; PRT; 100 AA.
AC P53312;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm histone P2 precursor (Protamine P2).
GN Pm2.
OS Alouatta seniculus (Howler monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;
OC Alouatta.
OX NCBI_TaxID=95503;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93292525; PubMed=8513810;
RA Retief J.D., Dixon G.H.;
RT "Evolution of pro-protamine P2 genes in primates."
RN Eur. J. Biochem. 214:609-615(1993).
[2]
RP ERRATUM.
RX MEDLINE=94109373; PubMed=8281927;
RA Retief J.D., Dixon G.H.;
RL Eur. J. Biochem. 218:1095-1095(1993).
-----
CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```
CC- between the Swiss Institute of Bioinformatics and the EMBL outstation
CC- the European Bioinformatics Institute. There are no restrictions on its
CC- use by non-profit institutions as long as its content is in no way
CC- modified and this statement is not removed. Usage by and for commercial
CC- entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC- or send an email to license@isb-sib.ch).
CC-----
DR EMBL; X71335; CAA50475.1; -.
DR PIR; S33338; S33338.
DR InterPro; IPR000492; Protamine_P2.
DR Pfam; PF00841; protamine_P2; 1.
KW Chromosomal protein, Nucleosome core, Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
FT PROPEP 1 ?
FT CHAIN 1 ?
SQ SEQUENCE 100 AA; 12560 MW; CBF579527B95CA84 CRC64;

Query Match 27.8%; Score 60; DB 1; Length 100;
Best Local Similarity 50.0%; Pred. No. 0.49;
Matches 19; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

Qy 11 RRVRRVVRVVRRVVRRVVRRVVRRVVRRVVRRVVRR 48
Db 62 RLRYRVARHORRCSCRCRCYRRNRNRCGCTRRRTCR 99

RESULT 4
HSP3_HORSE STANDARD; PRT; 58 AA.
AC PL543;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm histone P2B (ST2B).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mamalia; Euteria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE.
RX MEDLINE=90304188; PubMed=2364093;
RA Piiponen A., Valtanen P., Linnala-Kankkunen A., Heiskanen M.-L.,
RA Maenpaa P.K.;
RT "Primary structures of two protamine 2 variants (St2a and St2b) from
RL stallion spermatozoa.";
RN Blochim. Biophys. Acta 1039:177-180(1990).
RN [2]
RP SEQUENCE OF 1-25.
RX MEDLINE=89171259; PubMed=2924903;
RA Piiponen A., Linnala-Kankkunen A., Maenpaa P.K.;
RT "Comparison of partial amino acid sequences of two protamine 2
RT variants from stallion sperm. Structural evidence that the variants
RT are products of different genes.";
RN FEBS Lett. 244:199-202(1989).
RL
CC -1- FUNCTION: PROTAIINS SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.
DR PIR; S02787; S02787.
DR PIR; S10755; S10755.
KW Chromosomal protein, Nucleosome core, Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
SQ SEQUENCE 58 AA; 7979 MW; 8D1F3E09A873179 CRC64;

Query Match 27.3%; Score 59; DB 1; Length 58;
Best Local Similarity 50.0%; Pred. No. 0.38;
Matches 22; Conservative 2; Mismatches 18; Indels 2; Gaps 1;
```


CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-160

Query Match 32.6%; Score 62; DB 1; Length 28;
Best Local Similarity 42.9%; Pred. No. 0.03;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 6 RVRVRRVRRVRRVRRVRRVRRVRRV 33
DB 1 RVRVRRVRRVRRVRRVRRVRRVRRV 28

RESULT 8
US-08-932-682-160
Sequence 160, Application US/08932682
Patent No. 5945507
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mieczner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTS, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,682
FILING DATE: 18-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/786,748
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300

REFERENCE/DOCKET NUMBER: AP30421-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-705-5000
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5945507e
US-08-932-682-160

Query Match 32.6%; Score 62; DB 2; Length 28;
Best Local Similarity 42.9%; Pred. No. 0.03;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 6 RVRVRRVRRVRRVRRVRRVRRVRRV 33
DB 1 RVRVRRVRRVRRVRRVRRVRRVRRV 28

RESULT 9
US-08-231-730A-23
Sequence 23, Application US/08231730A
Patent No. 5561107
GENERAL INFORMATION:
APPLICANT: JULIAN, GORDON R.
TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST AN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: STEVEN J. HULTQUIST
STREET: 200 PARK DRIVE, SUITE 210
CITY: P.O. BOX 14329
STATE: RESEARCH TRIANGLE PARK
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: APPLE MACINTOSH
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231,730A
FILING DATE: 04-20-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-08-94
APPLICATION NUMBER: 08/039,620
FILING DATE: 06-04-93
APPLICATION NUMBER: 08/148,491
FILING DATE: 11-08-93
APPLICATION NUMBER: 08/148,889
FILING DATE: 11-08-93
ATTORNEY/AGENT INFORMATION:
NAME: HULTQUIST, STEVEN J.
REGISTRATION NUMBER: 28021
REFERENCE/DOCKET NUMBER: 4013-106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 37
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE

PRICE APPLICATION DATA:
APPLICATION NUMBER: 08/079,512
FILING DATE: 06/18/93
ATTORNEY/AGENT INFORMATION:
NAME: Roth, Michael J.
REGISTRATION NUMBER: 29,342
REFERENCE/DOCKET NUMBER: 0233 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 245-3594
TELEFAX: (515) 245-3634
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-179-632-9

Query Match 33.7%; Score 64; DB 1; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.019;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 12 RRVVRRVRRVRRVRRVRRVRRV 39
Db 1 RRIYRAIRHPRIRIGWLRIGRIERY 28

RESULT 5
US-08-440-174A-9
Sequence 9, Application US/08440174A
Patent No. 5717061
GENERAL INFORMATION:
APPLICANT: Rao, Gururaj A.
APPLICANT: Zhong, Lingxin
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: 7100 N.W. 62nd Avenue
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,174A
FILING DATE: 12-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/079,512
FILING DATE: 18-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bodrowicz, Donna
REGISTRATION NUMBER: 32,196
REFERENCE/DOCKET NUMBER: 0234R2D-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4896
TELEFAX: (515) 334-6883
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-174A-9

Query Match 33.7%; Score 64; DB 1; Length 31;

Best Local Similarity 42.9%; Pred. No. 0.019;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
Qy 12 RRVVRRVRRVRRVRRVRRVRRV 39
Db 1 RRIYRAIRHPRIRIGWLRIGRIERY 28

RESULT 6
PCT-US95-00062-9
Sequence 9, Application PC/TUS9500062
GENERAL INFORMATION:
APPLICANT: Pioneer Hi-Bred International, Inc.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/Microsoft Windows
SOFTWARE: Microsoft Windows Notepad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00062
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Yates, Michael E.; Sweeney, Patricia A.;
NAME: Roth, Michael J.; & Simon, Soma G.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 234R2-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00062-9

Query Match 33.7%; Score 64; DB 5; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.019;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 12 RRVVRRVRRVRRVRRVRRVRRV 39
Db 1 RRIYRAIRHPRIRIGWLRIGRIERY 28

RESULT 7
US-08-786-748A-160
Sequence 160, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mietzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza


```
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-058-12

Query Match      75.8%; Score 144; DB 9; Length 48;
Best Local Similarity 89.5%; Pred. No. 1.1e-11;
Matches 34; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 38
Db 7 RRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 44

RESULT 13
US-09-785-059-11
; Sequence 11, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A3577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-11

Query Match      70.5%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 1.3e-10;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 42
Db 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 14
US-10-079-075-11
; Sequence 11, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-11

Query Match      70.5%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 1.3e-10;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 42
Db 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 15
US-09-785-058-11
; Sequence 11, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-11

Query Match      70.5%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 1.3e-10;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 42
Db 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

Search completed: June 9, 2003, 12:34:09
Job time : 22.3404 secs
```

FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-6

Query Match 85.3%; Score 162; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.7e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 42
1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 8

US-10-079-075-6
Sequence 6, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-6

Query Match 85.3%; Score 162; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.7e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 42
Db 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 9

US-09-785-058-6
Sequence 6, Application US/09785058
Publication No. US2003003627A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001 / 072396.0222
CURRENT APPLICATION NUMBER: US/09/785.058
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-6

Query Match 85.3%; Score 162; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 4.7e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 7 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 42
1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 10

US-09-785-059-12
Sequence 12, Application US/09785059
Patent No. US20020169279A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-059-12

Query Match 75.8%; Score 144; DB 9; Length 48;
Best Local Similarity 89.5%; Pred. No. 1.1e-11;
Matches 34; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 38
Db 7 RRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 44

RESULT 11

US-10-079-075-12
Sequence 12, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: artificial peptides derived from HIV-1
US-10-079-075-12

Query Match 75.8%; Score 144; DB 9; Length 48;
Best Local Similarity 89.5%; Pred. No. 1.1e-11;
Matches 34; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 38
Db 7 RRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 44

RESULT 12

US-09-785-058-12
Sequence 12, Application US/09785058
Publication No. US2003003627A1

PN WO603519-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US09339.
XX
PR 22-JUL-1994; 94US-0279472.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX (USDA) US SEC OF AGRIC.
XX
PI Belknap W, Garbarino J, Jaynes J;
XX
DR WPI; 1996-117061/12.
XX
PT New fusion protein of ubiquitin and a lytic peptide - for treating
XX infections and neoplasia, healing wounds, etc. also related nucleic
PT acid, vectors, and transformed cells
XX
PS Claim 5; Page 19; 112pp; English.
XX
XX AAR92372-R92462 are lytic peptides used to create ubiquitin-lytic
CC peptide fusion proteins in which the ubiquitin polypeptide is linked
CC at its 3'-terminus to the lytic peptide. The lytic peptides are pref-
CC selected from either the cecropins, defensins, sarcotoxins, melittin
CC and magainins. The fusion proteins (FPs) are useful for treating (in
CC protozoal, bacterial, fungal and viral infections and neoplasia (in
CC plants and animals) in the same way as the FP alone, they also
CC promote wound healing. FPs produced in bacteria may be cleaved in
CC vitro by ubiquitin hydrolases to recover the active lytic peptide.
CC FPs produced in eukaryotic cells are cleaved by endogenous enzymes
CC to yield lytic peptide. Recombinant DNA encoding the FPs have
CC greater stability in bacteria than DNA encoding the lytic peptide
CC only.
XX
SQ Sequence 37 AA;
XX
Query Match 31.6%; Score 60; DB 17; Length 37;
Best Local Similarity 26.5%; Pred. No. 0.49;
Matches 9; Conservative 15; Mismatches 10; Indels 0; Gaps 0;
OY 1 RRVRRRRRVRRRVRRRVRRRVRRRVRR 34
DB 3 KKFVKVAKVAKVAKVAKVAKVAKVAK 36
RESULT 14
AAR90748
ID AAR90748 standard; peptide; 37 AA.
XX
AC AAR90748;
XX
DT 17-SEP-1996 (first entry)
XX
DE Synthetic lytic peptide #78.
XX
XX Lytic peptide; ubiquitin; synthetic analogue; cell membrane; cell lysis;
KW microbial pathogen; disease-resistant plant; bacterial infection; fungus;
KM protozoa; virus; neoplasia; fusion protein; hydrolase.
XX
OS Synthetic.
XX
XX WO603522-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US09338.
XX
PR 22-JUL-1994; 94US-0279472.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
PI Jaynes J;
XX

XX WPI; 1996-117064/12.
DR
XX
XX Lytic peptide(s), useful for developing disease-resistant plants -
PT can be expressed as fusion protein with ubiquitin for stable prodn.
PT in bacterial host cells
XX
PS Example 1; Page 19; 111pp; English.
XX
XX AAR89967-R90021 and AAR90726-R90763 represent synthetic analogues of
CC naturally occurring lytic peptides. Lytic peptides destroy prokaryotic
CC and other non-host cells by disrupting the cell membrane and promoting
CC cell lysis. Synthetic lytic peptide analogues have similar or higher
CC levels of lytic activity for many different types of cells, compared to
CC naturally occurring forms. The concentration of the synthetic analogue
CC required to lyse microbial pathogens does not lyse normal mammalian
CC cells. The lytic peptides can be expressed in plants to allow for the
CC development of disease-resistant plants. The peptides are useful in
CC promoting wound healing and combating bacterial infections in plants.
CC The lytic peptides can also be used for combating protozoal, fungal,
CC viral or bacterial infections or neoplasias in mammals and plants.
CC Lytic peptide-ubiquitin fusion proteins are suitable for production in
CC bacterial hosts. Bacteria lack the hydrolase which cleaves the peptide
CC from ubiquitin, and therefore the active (and cytotoxic) lytic peptide
CC will not be released in the host cells. The recombinantly produced
CC lytic peptide can be retrieved from the fusion protein by cleavage in
CC vitro.
XX
SQ Sequence 37 AA;
XX
Query Match 31.6%; Score 60; DB 17; Length 37;
Best Local Similarity 26.5%; Pred. No. 0.49;
Matches 9; Conservative 15; Mismatches 10; Indels 0; Gaps 0;
OY 1 RRVRRRRRVRRRVRRRVRRRVRRRVRR 34
DB 3 KKFVKVAKVAKVAKVAKVAKVAKVAK 36
RESULT 15
AAM39765
ID AAM39765 standard; peptide; 37 AA.
XX
AC AAM39765;
XX
DT 20-MAY-1998 (first entry)
XX
DE Synthetic lysine-rich lytic peptide 16.
XX
XX Lytic peptide; lysine-rich; proteolytic digestion; methylation; trypsin;
KW protection; amphipathic alpha-helix; beta-pleated sheet; treatment;
KM infection; viral; bacterial; yeast; fungal; protozoan; cancer.
XX
OS Synthetic.
XX
XX US5717064-A.
XX
PD 10-FEB-1998.
XX
PF 24-APR-1995; 95US-0427001.
XX
PR 24-APR-1995; 95US-0427001.
XX
PR 18-NOV-1993; 93US-0148889.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
XX Jaynes JM, Julian GR;
PI WPI; 1998-158370/14.
DR
XX
XX Cytolytic peptide analogues - with methylated lysine residues to
PT increase protease resistance
PT
XX

[illegible]

DE	Amphipathic peptide #23, for treating a pulmonary disease state.
XX	
XX	Amphipathic peptide; pulmonary disease; resistant; bacteria;
KW	proteolytic digestion; methylation; glyoxylation; cystic fibrosis;
CF; neoplasia; pneumonia; bronchitis; lytic activity; lysis.	
XX	
OS	Synthetic.
XX	
FH	Key
FT	Modified-site 1 Location/Qualifiers
FT	/note= "Opt. methylated or glyoxylated"
FT	1..37
FT	/note= "Lys residues are opt. methylated and/or Arg residues are glyoxylated"
FT	
PN	WO9428921-A.
XX	
PD	22-DEC-1994.
XX	
PF	02-JUN-1994; 94WO-US06176.
XX	
PR	04-JUN-1993; 93US-0039620.
XX	
PA	(DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX	
P1	Jaynes JM, Julian GR;
DR	WPI; 1995-036106/05.
XX	
PT	Treatment of pulmonary disease states such as cystic fibrosis - by admin. of a non-naturally occurring amphipathic peptide
PS	Claim 21; Page 49; 54pp; English.
XX	
CC	The sequences given in AAR64770-806 and AAR76077 represent non-naturally occurring amphipathic peptides which may be used for treating
CC	pulmonary disease states. These peptides have enhanced resistance
CC	to proteolytic digestion due to methylation of the epsilon-amino
CC	group of lys residues or glyoxylation of the side chains of Arg
CC	residues, and methylation or glyoxylation of the N-terminal amino
CC	acid. The peptides are prepared by standard methods of solid phase
CC	synthesis and may be used in the treatment of cystic fibrosis (CF),
CC	neoplasia, pneumonia, bronchitis, etc.. The peptides pref. have a
CC	lytic activity, thereby lysing pathogenic bacteria, virally infected
CC	cells and transformed cells as well as treating the epithelial cell
CC	defect of CF.
XX	
SO	Sequence 37 AA;
Query Match	31.6%; Score 60; DB 16; Length 37;
Best Local Similarity	26.5%; Pred. No. 0.49;
Matches 9; Conservative 15; Mismatches 10; Indels 0; Gaps 0;	
DQ	1 RRVVRRVVRRVVRRVVRRVVRRVVRRVR 34 :: :: :: :: :: :: :: :: :: :: :
Db	3 KKFVKVAKVAKKVAKKVAKKVAKKVAKVAK 36
RESULT 13	
AAR92394	
ID	AAR92394 standard; peptide; 37 AA.
AC	AAR92394;
DT	17-SEP-1996 (first entry)
XX	
DE	Lytic peptide used in ubiquitin-lytic peptide fusion protein.
XX	
KW	Ubiquitin; fusion protein; lysis; infection; neoplasia; wound healing;
XX	stability; reduced toxicity.
OS	Synthetic.
XX	

XX Disclosure; Column 21; 62pp; English.
XX
XX This sequence represents an antimicrobial peptide of the invention, and
CC is an analogue of the peptide LRP1 (see AAY32549). The peptides can be
CC used for treating infections caused by *Staphylococcus aureus*,
CC methicillin resistant *S. aureus*, *Pseudomonas aeruginosa*, *Enterococcus*
CC *faecalis*, *S. marcescens*, *Bacterioides coli*, fungi, protozoa and viruses in
CC a mammalian host. They can be used to inhibit growth of diverse
CC microorganisms such as bacteria, fungi, protozoa and DNA and RNA viruses
CC and can be used in tissue culture to inhibit unwanted microbial growth,
CC particularly for the production of recombinant proteins or vectors for
CC gene therapy. They can also be used in preventing infections through the
CC sterilisation of wounds prior to suture and to sterilise surgical
CC instruments. The unique structure of these antimicrobial peptides
CC imparts high potency while selectivity is maintained, they are
CC moderately haemolytic but only lyse red blood cells at high
CC concentrations unlike melittin, a peptide extracted from bee venom, which
CC is highly active against bacteria and lyses red blood cells showing
CC little selectivity. The peptides target a membrane structure which makes
CC it more difficult for a microorganism to develop a mechanism of
CC resistance against this type of antibiotic. Their small size makes them
CC relatively simple to prepare by standard synthetic peptide chemistry.
XX
SQ Sequence 28 AA;
Query Match 32.6%; Score 62; DB 20; Length 28;
Best Local Similarity 42.9%; Pred. No. 0.21;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
QY 6 RRVRRVRRVRRVRRVRRVRRVRRV 33
DB 1 RVIRVVGACRAIRHPRIRIOGLRLT 28
RESULT 9
AAR84150
ID AAR84150 standard; Peptide; 37 AA.
XX
AC AAR84150;
XX
DT 06-JUN-1996 (first entry)
XX
DE Peptide enhancer of fibroblast and keratinocyte proliferation.
XX
KM amphipathic peptide; enhance; fibroblast; keratinocyte; proliferation;
KM wound healing; defensin; antimicrobial.
XX
OS Synthetic.
XX
PN WO9528832-A1.
XX
PD 02-NOV-1995.
XX
PF 19-APR-1995; 95WO-US04718.
XX
PR 20-APR-1994; 94US-0231730.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
PI Jaynes JM, Julian GR;
XX
DR WPI; 1995-382791/49.
XX
PT Use of amphipathic peptide(s) - for enhancing fibroblast and
PT keratinocyte proliferation to promote wound healing in a mammal
XX
XX Claim 7; Page 49; 64pp; English.
XX
CC AAR84128-73 are amphipathic peptides which are able to stimulate the
CC proliferative growth of fibroblasts and epithelial cells such as
CC keratinocytes, hence enhancing wound healing in mammalian subjects.
CC The peptides concomitantly have antimicrobial efficacy, against

CC microbial species including those which cause or otherwise mediate
CC sepsis and wound infection.
XX
SQ Sequence 37 AA;
Query Match 31.6%; Score 60; DB 16; Length 37;
Best Local Similarity 26.5%; Pred. No. 0.49;
Matches 9; Conservative 15; Mismatches 10; Indels 0; Gaps 0;
QY 1 RRVRRVRRVRRVRRVRRVRRVRRVRRV 34
DB 3 KKKVKKVAVAKKAVAKKAVAKKAVAKK 36
RESULT 10
AAR77064
ID AAR77064 standard; peptide; 37 AA.
XX
AC AAR77064;
XX
DT 20-MAY-1996 (first entry)
XX
DE Synthetic anti-neoplastic lytic peptide.
XX
KM Anti-cancer; lysis; amphipathic; neoplasia; tumour; cystic fibrosis;
KM bronchopulmonary; viral; virus; analogue; magainin; cecropin;
KM melittin; defensin.
XX
OS Synthetic.
XX
PN WO9527497-A1.
XX
PD 19-OCT-1995.
XX
PF 06-APR-1995; 95WO-US04335.
XX
PR 08-APR-1994; 94US-0225476.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
PI Jaynes JM, Julian GR;
XX
DR WPI; 1995-366226/47.
XX
PT Method of combating mammalian neoplasias and other disease states -
PT by delivering non-naturally occurring, non-onco-cytologically
PT proliferative lytic peptide
XX
PS Claim 7; Page 47; 63pp; English.
XX
CC AAR77042-R77081 are synthetic, amphipathic, lytic peptide analogues of
CC melittin, cecropin, magainin and defensin peptides. The peptides are
CC between 23 and 39 residues long, are amphipathic, carry an overall
CC positive charge and have anti-neoplastic activity. The peptides are
CC specifically useful for the lysis of cancer cells. Normal mammalian
CC cells are resistant to lysis due to their highly organised
CC cytoskeleton, cancerous cells however possess an interior and
CC structurally compromised cytoskeleton which when acted upon by lytic
CC peptides will cause cell lysis. This allows the lytic peptides to be
CC used for in vivo treatment of cancers. The peptides are esp. useful
CC for the treatment of female mammalian cancers e.g. breast, ovarian,
CC uterine and cervical cancers. The peptides can however be used to
CC treat most forms of cancer, cystic fibrosis, pneumonia, bronchitis,
CC and bronchopulmonary viral and microbial infections.
XX
SQ Sequence 37 AA;
Query Match 31.6%; Score 60; DB 16; Length 37;
Best Local Similarity 26.5%; Pred. No. 0.49;
Matches 9; Conservative 15; Mismatches 10; Indels 0; Gaps 0;
QY 1 RRVRRVRRVRRVRRVRRVRRVRRVRRV 34
DB 3 KKKVKKVAVAKKAVAKKAVAKKAVAKK 36

KM Antimicrobial; antifungal; pathogen; plant; amphipathic;
XX broad spectrum.
OS Synthetic.
XX WO9518855-A2.
XX
PD 13-JUL-1995.
XX
PF 06-JAN-1995; 95MO-US00062.
XX
PR 07-JAN-1994; 94US-0179632.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Rao AG, Zhong L;
XX
DR WPI; 1995-255059/33.
XX
PT New antimicrobial amphipathic polypeptide(s) and related nucleic
PT acids - for clinical use or esp. to increase resistance of plants
PT to fungal pathogens.
XX
XX
PS Claim 1; Page 17; 24pp; English.
XX
CC AAR80727-R80748 are new amphipathic polypeptides. They have a broad
CC spectrum of antimicrobial and antifungal activity. They can be used
CC to treat or prevent infection in humans and animals or applied to
CC plants as sprays, creams, dust, etc. The DNA encoding these
CC peptides can also be incorporated into susceptible plants via the
CC use of a non-phytotoxic vehicle adapted for systemic administration.
CC This process imparts resistance to plant pathogens esp. fungi (e.g.
CC Fusarium graminearum, F.moliforme, Aspergillus flavus, Alternaria
CC longipes, Colletotrichum graminicola, Phytophthora megasperme,
CC Sclerotinia sclerotiorum). The peptides are esp. useful in
CC transformed plants such as maize, sorghum, wheat, soya, alfalfa,
CC rapeseed, sunflower, tobacco or tomato.
XX
SQ Sequence 31 AA;

Query Match 33.7%; Score 64; DB 16; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.13;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 12 RRVRRVRRVRRVRRVRRVRRVRRV 39
DB 1 RRIYRAIRHLPRIIRIGWLRIGRIERY 28

RESULT 7
AAW47769 standard; peptide; 28 AA.
ID AAW47769
XX
XX AAW47769;
AC
XX
XX 26-MAY-1998 (first entry)
DT
XX
XX Antimicrobial peptide LRP1 analogue.
DE
XX Antimicrobial; transmembrane protein; TM; lentivirus lytic peptide;
KW LRP; amphipathic; antibacterial; antifungal; antiviral; antiprotocool.
XX
XX Synthetic.
OS Human immunodeficiency virus.
XX
XX US5714577-A.
XX
XX
XX 03-FEB-1998.
PD
XX
XX 24-JAN-1997; 97US-0786748.
PF
XX
XX 26-JAN-1996; 96US-0010634.
PR
XX 24-JAN-1997; 97US-0786748.
PT

XX
XX (UYPI-) UNIV PITTSBURGH.
XX
XX Mietzner TA, Montelaro RC, Tencza SB;
PI
XX WPI; 1998-158352/14.
DR
XX
XX Retroviral TM peptides - useful as antibacterial agents
PT
XX
PS Disclosure; Column 19; 59pp; English.
XX
XX The invention relates to new antimicrobial peptides which correspond to
XX amino acid sequences in the transmembrane proteins of lentiviruses, in
XX particular HIV and SIV. These peptides comprise arginine rich sequences
XX which, when modelled for secondary structure, display high
XX amphipathicity and hydrophobic moment. Also disclosed are structural
XX and functional analogues and homologues of these peptides which also
XX display antimicrobial activity. The peptides are highly inhibitory to
XX microorganisms (bacteria, fungi, viruses and protozoa) but significantly
XX less toxic to red blood cells and other normal mammalian cells. Activity
XX is demonstrated against Gram positive and negative bacteria including
XX Pseudomonas aeruginosa, Staphylococcus aureus, Enterococcus faecalis and
XX Serratia marcescens.
XX The present sequence is one of 169 disclosed specific examples of
XX the new peptides. It is an analogue of the peptide designated LRP1
XX (see AAW47614) which is a peptide from the transmembrane protein (gp41)
XX of HIV strain HXB2R.
SQ Sequence 28 AA;

Query Match 32.6%; Score 62; DB 19; Length 28;
Best Local Similarity 42.9%; Pred. No. 0.21;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 6 RRVRRVRRVRRVRRVRRVRRVRRV 33
DB 1 RVIKRVQACRAIRHLPRIIRIGLRIL 28

RESULT 8
AAV32703 standard; peptide; 28 AA.
ID AAV32703
XX
XX AAV32703;
AC
XX
XX 21-OCT-1999 (first entry)
DT
XX
XX Antimicrobial peptide LRP1 analogue.
DE
XX Antimicrobial peptide; LRP1; SLP-1; LRP2; SLP2A; SLP2B; ELP; infection;
KW growth inhibitor; microorganism; virus; gene therapy; vector production;
KW sterilisation.
XX
XX Synthetic.
OS Human immunodeficiency virus type 1.
XX
XX US5945507-A.
XX
XX
XX 31-AUG-1999.
PD
XX
XX 18-SEP-1997; 97US-0932682.
PF
XX
XX 26-JAN-1996; 96US-0010634.
PR
XX 24-JAN-1997; 97US-0786748.
PR
XX 18-SEP-1997; 97US-0932682.
XX
XX (UYPI-) UNIV PITTSBURGH.
XX
XX
XX Mietzner TA, Montelaro RC, Tencza SB;
PI
XX WPI; 1999-508189/42.
DR
XX
XX Antimicrobial peptides useful for treating microbial infections
PT

[illegible]

SO	Sequence	18 AA;
Oy	Query Match	35.3%; Score 67; DB 22; Length 18;
	Best Local Similarity	47.1%; Pred. No. 0.034;
	Matches	8; Conservative 9; Mismatches 0; Indels 0; Gaps 0;
Db	9 RVRVRRVVRRVVRRRV 25 :: ::: :: : 2 RIIRRIIRIIRIIRRI 18	
RESULT 5		
ID	AAR60065 standard; peptide; 31 AA.	
AC	AAR60065;	
DT	16-FEB-1995 (first entry)	
DE	Antimicrobial peptide.	
KW	Amphipathic peptide; alpha-helix; lytic peptide; antifungal;	
RK	antimicrobial; fungus resistance; disease resistance;	
OS	crop protection.	
XX	Synthetic.	
XX	MO9415961-A.	
PN	21-JUL-1994.	
PD	12-JAN-1994; 94WO-US00383.	
PF	13-JAN-1993; 93US-0003884.	
PR	(PION-) PIONEER HI-BRED INT INC.	
PA	Rao AG, Zhong L;	
PI	WPI; 1994-249137/30.	
DR	Synthetic polypeptide(s) and the nucleic acid encoding them -	
PT	exhibits amphipathic alpha-helices and provide cell-expressable	
FT	antimicrobial activity	
XX	Disclosure; Page 19; 25pp; English.	
PS	The synthetic amphipathic alpha-helical lytic peptides given in	
CC	AAR60057-71 were designed to provide antifungal or antimicrobial	
CC	activity when expressed in monocot or dicot plants. They also have	
CC	veterinary and medical applications.	
XX		
SO	Sequence	31 AA;
Oy	Query Match	33.7%; Score 64; DB 15; Length 31;
	Best Local Similarity	42.9%; Pred. No. 0.13;
	Matches	12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
Db	12 RRVRVRRVRRVRRVRRVRRV 39 :: ::: :: : 1 RRIYRAIRHIPPRIKRWLRIGRRIERY 28	
RESULT 6		
ID	AAR80735 standard; peptide; 31 AA.	
AC	AAR80735;	
DT	28-FEB-1996 (first entry)	
DE	Synthetic antimicrobial/antifungal polypeptide.	
XX		


```

DR  WP1, 1995-276981/37.
XX
XX  Complex of nucleic acid and oligopeptide with sec. structure - and
PT  transfer vectors contg. them, useful for efficient transfer of
PT  nucleic acid to cells in gene therapy.
XX
XX  Claim 6; Page 16; 20pp; French.
XX
XX  The present peptide corresponds to a generic formula for a cationic
CC  oligopeptide; the formula is (b-1-1-b)n, where b is a hydrophobic amino
CC  acid, 1 is a hydrophilic amino acid and n is at least 4. In this case,
CC  where b is leu, 1 is Arg and n = 10, the oligopeptide forms an alpha-
CC  helix which forms a stable complex with a nucleic acid. The complex
CC  is suitable for transferring nucleic acid, esp. in gene therapy.
XX
XX  Sequence 40 AA;
SQ
XX
XX  Query Match 40.3%; Score 76.5; DB 16; Length 40;
XX  Best Local Similarity 46.2%; Pred. No. 0.0053;
XX  Matches 18; Conservative 18; Mismatches 0; Indels 3; Gaps 3
Qy 7 VRRVRRVV-RVRRRRVR-VRRVRRVV-RVRRRRVR 42
Db 1 LRRLLRRLLRRLLRRLLRRLLRRLLRRLLRRLLRR 39
XX
XX  RESULT 2
XX  AAM06684
XX  AAM06684 standard; peptide; 39 AA.
XX
XX  AAM06684;
XX
XX  05-AUG-1997 (first entry)
DE
XX  Protamine-like peptide analogue [+18RGD].
XX
XX  Coagulation; anticoagulant; heparin; platelet aggregation; cell adhesion;
KM  positively charged cluster; arginine; polycationic; decrease;
KM  n-protamine; salmine protamine; protamine sulphate; salmon sperm.
XX
XX  Synthetic.
XX
XX  Key Location/Qualifiers
XX  FT Modified-site 1
XX  FT Region 3..30
XX  FT /note= "acetylated"
XX  FT /label= repeat region
XX  FT Binding-site /note= "4 tandem repeats of (Arg)2(Ala)2Ala motif"
XX  FT 34..36
XX  FT /label= cell_adhesion_motif
XX  FT Modified-site 39
XX  FT /note= "amidated"
XX
XX  WO9635444-A1.
XX
XX  14-NOV-1996.
XX
XX  08-MAY-1996; 96WO-US06567.
XX
XX  08-MAY-1995; 95US-0436703.
XX
XX  (UNMI ) UNIV MICHIGAN.
XX
XX  Andrews PC, Stanley JC, Wakefield TW;
XX
XX  WPI, 1997-011697/01.
XX
XX  Peptide reversing the anticoagulant effects of heparin - is based on
XX  protamine but has fewer positive charges for reduced toxicity
XX
XX  Claim 31; Page 31; 42pp; English.
XX
XX  Protamine sulphate (also called n-protamine or salmine protamine) is
CC

```

[illegible]


```
RN [1] SEQUENCE FROM N.A.
RP STRAIN=LV39;
RX MEDLINE=94187808; PubMed=8139626;
RA Fasel N.J., Robyr D.C., Manuel J., Glaser T.A.;
RT "Identification of a histone H1-like gene expressed in Leishmania
   major.";
RL Mol. Biochem. Parasitol. 62:321-323(1993).
DR EMBL, U01031; AAA18635.1; -.
SQ SEQUENCE 105 AA; 10909 MW; D8C32835131ACE38 CRC64;

Query Match
Best Local Similarity 31.6%; Score 60; DB 5; Length 105;
Matches 12; Conservative 18; Mismatches 7; Indels 2; Gaps 2;

QY 1 RRVRR-VRVRR-VRVRRVRVRRVRVRRVRVRR 37
   ::|||: ::|||: ::|||: ::|||: ::|||:
DB 54 KKVKKPAKKVKKPAKKVKKVKKVKKVKKVKKV 92

RESULT 13
O9TV18 PRELIMINARY; PRT; 105 AA.
ID O9TV18;
AC O9TV18;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE Histone H1.
GN SW3.1 OR SW3.0.
OS Leishmania major.
OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=FRIEDLIN;
RA Fasel N.J.;
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2] SEQUENCE FROM N.A.
RP STRAIN=FRIEDLIN;
RA Fasel N.J., Robyr D., Manuel J., Glaser T.A.;
RT "Identification of a histone H1-like gene expressed in Leishmania
   major.";
RL Mol. Biochem. Parasitol. 62:321-324(1994).
RN [3] SEQUENCE FROM N.A.
RP STRAIN=FRIEDLIN;
RX MEDLINE=97237563; PubMed=9084041;
RA Noll T., Desponds C., Jacques R., Belli S., Fasel N.J.;
RT "T. Noll, C. Desponds, R. Jacques, S. Belli and N. J. Fasel. Histone
   H1 expression varies during Leishmania major development.";
RL Mol. Biochem. Parasitol. 84:215-227(1997).
DR EMBL, AJ223861; CAA11592.1; -.
DR EMBL, AJ223860; CAA11591.1; -.
SQ SEQUENCE 105 AA; 10819 MW; D9729835131BCE38 CRC64;

Query Match
Best Local Similarity 31.6%; Score 60; DB 5; Length 105;
Matches 12; Conservative 18; Mismatches 7; Indels 2; Gaps 2;

QY 1 RRVRR-VRVRR-VRVRRVRVRRVRVRRVRVRR 37
   ::|||: ::|||: ::|||: ::|||: ::|||:
DB 54 KKVKKPAKKVKKPAKKVKKVKKVKKVKKVKKV 92

RESULT 14
O91D00 PRELIMINARY; PRT; 755 AA.
ID O91D00;
AC O91D00;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ORF1.
```

```
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1] SEQUENCE FROM N.A.
RX MEDLINE=2148921; PubMed=11601907;
RA Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
   Yoshikawa A.;
RT "Heterogeneous distribution of TT virus of distinct genotypes in
   multiple tissues from infected humans.";
RL Virology 288:358-368(2001).
DR EMBL, AB060593; BAB63904.1; -.
DR InterPro, IPR004219; TTVirus_unk.
DR Pfam, PF02956; TT_ORF1_1.
SQ SEQUENCE 755 AA; 89785 MW; 3ABC076D93F1F8FA CRC64;

Query Match
Best Local Similarity 31.6%; Score 60; DB 12; Length 755;
Matches 21; Conservative 3; Mismatches 12; Indels 4; Gaps 2;

QY 5 RRVRRV--RRVRRVRVRRVRVRRVRVRRVRVRR 42
   ||||: ||||: ||||: ||||: ||||:
DB 33 RRPRLARGRRRRRTVRR--RRVRLRRRGWTRRRVLR 70

RESULT 15
O90CG7 PRELIMINARY; PRT; 862 AA.
ID O90CG7;
AC O90CG7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=CM53379;
RX MEDLINE=21342568; PubMed=11448170;
RA Carr J.K., Torimiro J.N., Wolfe N.D., Etzel M.N., Kim B.,
   Sanders-Buell E., Jagodzinski L.L., Gotte D., Burke D.S., Bix D.L.,
   McCutchan F.E.;
RT "The AG recombinant IBNG and novel strains of group M HIV-1 are common
   in Cameroon.";
RL Virology 286:168-181(2001).
DR EMBL, AF377959; AAK59217.1; -.
DR InterPro, IPR000328; Env_Gp41.
DR Pfam, PF00516; GP120; 1.
DR Pfam, PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 862 AA; 97328 MW; C93A1D72B809961 CRC64;

Query Match
Best Local Similarity 31.6%; Score 60; DB 15; Length 862;
Matches 11; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 6 RRVRRVRVRRVRVRRVRVRRVRVRRVRVRR 34
   ||||: ||||: ||||: ||||: ||||:
DB 834 RVIEIVRATRAIRNIPRIQGAERALQ 862

Search completed: June 9, 2003, 12:01:08
Job time : 39.5745 secs
```

OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Matsumoto A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.,
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AB003010; BABS3093.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 349 AA; 37473 MW; B7E34ECC39304 CRC64;
 QY
 Query Match 33.7%; Score 64; DB 16; Length 349;
 Best Local Similarity 48.4%; Pred. No. 4.3;
 Matches 15; Conservative 3; Mismatches 13; Indels 0; Gaps 0;
 Db 5 RRVRRVRRVRRVRRVRRVRRVRRVRRV 35
 106 RRIARGVGVRRVRRVRRVRRVRRVRRV 136
 RESULT 9
 Q91C22. PRELIMINARY; PRT; 759 AA.
 AC Q91C22;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ORF1.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21488921; PubMed=11601907;
 RA Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
 RA Yoshikawa A.,
 RT "Heterogeneous distribution of TT virus of distinct genotypes in
 multiple tissues from infected humans.";
 RL Virology 288:358-368(2001).
 DR EMBL: AB060595; BAB69912.1;
 DR InterPro: IPR004219; TVvirus_Unk.
 DR Pfam: PF02956; TT_ORF1.1.
 SQ SEQUENCE 759 AA; 89860 MW; SB79DFD71A37010D CRC64;
 QY
 Query Match 33.4%; Score 63.5; DB 12; Length 759;
 Best Local Similarity 51.0%; Pred. No. 10;
 Matches 25; Conservative 1; Mismatches 14; Indels 9; Gaps 3;
 Db 1 RRVRR-----VRRVRRVRRVRRVRRVRRVRRVRRVRR 42
 25 RRIARRRRRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 71
 RESULT 10
 Q8TXS5. PRELIMINARY; PRT; 428 AA.
 AC Q8TXS5;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Uncharacterized protein.
 GN MK0585.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polunin N.N.,
 RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Nale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malyshev A.G., Koonin E.V., Kozlovskiy S.A.,
 RT "The complete genome of hyperthermophilic Methanopyrus kandleri AV19
 and morphology of archaeal methanogens.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL: AB010352; AAM01800.1;
 KW Complete proteome.
 SQ SEQUENCE 428 AA; 48744 MW; 964D47CA264D3C CRC64;
 QY
 Query Match 33.2%; Score 63; DB 17; Length 428;
 Best Local Similarity 39.0%; Pred. No. 6.7;
 Matches 16; Conservative 12; Mismatches 9; Indels 4; Gaps 2;
 Db 1 RRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 39
 356 QKURREVERRYTGR--RLVRKIVRKIARVLRKEDFRVEGKI 394
 RESULT 11
 Q91PS3. PRELIMINARY; PRT; 742 AA.
 AC Q91PS3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ORF1.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KT-10F;
 RA Okamoto H.,
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=KT-10F;
 RX MEDLINE=21440417; PubMed=11556704;
 RA Muljono D.H., Nishizawa T., Tsuda F., Takahashi M., Okamoto H.,
 RT "Molecular epidemiology of TT virus (TTV) and characterization of two
 RT novel TTV genotypes in Indonesia.";
 RL Arch. Virol. 146:1249-1266(2001).
 DR EMBL: AB054648; BAB61611.1;
 DR InterPro: IPR004219; TVvirus_Unk.
 DR Pfam: PF02956; TT_ORF1.1.
 SQ SEQUENCE 742 AA; 86114 MW; A4E27AB09163DB5A CRC64;
 QY
 Query Match 33.2%; Score 63; DB 12; Length 742;
 Best Local Similarity 52.5%; Pred. No. 11;
 Matches 21; Conservative 2; Mismatches 13; Indels 4; Gaps 2;
 Db 5 RRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 42
 31 RRTTRVARR--RRVRLRRRRRRRGWARRRYLRARRRRVRR 68
 RESULT 12
 Q25305. PRELIMINARY; PRT; 105 AA.
 AC Q25305;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE Histone H-1 like protein.
 OS Leishmania major.
 OC Eukaryota; Euzoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;

ID	Q99AR5		PRELIMINARY;	PRT;	760 AA.
AC	Q99AR5;				
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	Orf1.				
OS	TT virus.				
OC	Viruses; ssDNA viruses; unclassified ssDNA viruses.				
OX	NCBI_TaxID=68887;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=TCHN-G1;				
RA	Luo K.-X., He H.-T., Liu D.-X., Liu Z.-H., Xao H., Jiang X.-J., Liang W.-F., Zhang L.;				
RT	"Novel variants related to TT virus wide distribution in China.";				
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF345521; AAK11696.1; "				
DR	InterPro; IPR004219; TTVirus_Unk.				
SQ	Pfam; PF02956; TT_ORF1.1.				
	SEQUENCE 760 AA; 90494 MW; 50BD115CCF55181A CRC64;				
Query Match		35.5%; Score 67.5; DB 12; Length 760;			
Best local Similarity		46.7%; Pred. NO. 3.6;			
Matches	21; Conservative	3; Mismatches 18; Indels 3; Gaps 1;			
CY	1 RRVRRRVRRVVRRVVRRVVRRVVRRVVRR--VVRVVRR 42				
DB	25 RLPTRTTRRAVAGLGRRPKTKVTTRRRRRRPRRTYRGWRERYIR 69				
RESULT 6					
O84529		PRELIMINARY;	PRT;	114 AA.	
ID	Q84529				
AC	O84529;				
DT	01-NOV-1996	(TREMBLrel. 01, Created)			
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	A209R protein.				
GN	A209R.				
OC	Parametium bursaria chlorella virus 1 (PBCV-1)				
OC	Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.				
OX	NCBI_TaxID=10506;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96187795; PubMed=8614977;				
RA	Lu Z., Li Y., Que O., Kutish G.F., Rock D.L., Van Etten J.L.;				
RT	"Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: map positions 88 to 182.";				
VL	Virology 216:102-123(1996).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20013326; PubMed=10544099;				
RA	Kaiser A., Volmert M., Tholl D., Graves M.V., Gurron J.R., Xing W., Liasec A.D., Nickerson K.W., Van Etten J.L.;				
RT	"Chlorella virus PBCV-1 encodes a functional homospesmidine synthase.";				
VL	Virology 263:254-262(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20478054; PubMed=11021991;				
RA	Sun L., Gurron J.R., Adams B.J., Graves M.V., Van Etten J.L.;				
RT	"Characterization of a beta-1,3-glucanase encoded by chlorella virus PBCV-1.";				
VL	Virology 276:27-36(2000).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Van Etten J.L.;				
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.				
NN	[5]				

RP	SEQUENCE FROM N.A.
RA	Van Etten J.L.;
RL	Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN	[6]
RP	SEQUENCE FROM N.A.
RA	Van Etten J.L.;
RL	Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN	[7]
RP	SEQUENCE FROM N.A.
RA	Graves M.V., Van Etten J.L.;
RL	Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN	[8]
RP	SEQUENCE FROM N.A.
RA	Graves M.V., Van Etten J.L.;
RL	Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN	[9]
RP	SEQUENCE FROM N.A.
RA	Gurion J.R., Graves M.V., Van Etten J.L.;
RL	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; U42580; AAC96577.1; -
SQ	SEQUENCE 114 AA; 14620 MW; 50E4B59A28366D3 CRC64;
OY	Query Match 33.7%; Score 64; DB 12; Length 114; Best Local Similarity 45.7%; Pred.No.1.5;
Dn	Matches 21; Conservative 7; Mismatches 12; Indels 6; Gaps 2;
RESULT 7	
ID	Q25562 PRELIMINARY; PRT; 302 AA.
AC	Q25562;
DT	01-NOV-1996 (TREMBlrel. 01, Created)
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE	ARK-rich protein (Fragment).
OS	Naegleria gruberii.
OC	Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX	NCHI_TaxID=5762;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=NEG-M;
RA	Clark C.G.;
RT	"Genome structure and evolution in the amoeba-flagellate protozoan
RL	Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; L139936; AAA62487.1; -
FT	NON TER 1
SQ	SEQUENCE 302 AA; 34677 MW; 9EA5DDA062F19A45 CRC64;
OY	Query Match 33.7%; Score 64; DB 5; Length 302; Best Local Similarity 27.1%; Pred.No.3.7;
Dn	Matches 13; Conservative 14; Mismatches 15; Indels 6; Gaps 1;
RESULT 8	
ID	Q987V7 PRELIMINARY; PRT; 349 AA.
AC	Q987V7;
DT	01-OCT-2001 (TREMBlrel. 18, Created)
DT	01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE	Hypothetical protein ml16891.
GN	ML16891.
OS	Rhizobium loti (Mesorhizobium loti).
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

RC STRAIN=M8R-1;
 RA MEDLINE=99119503; PubMed=9918888;
 RA Reddy P.S., Chen Y., Idamakanti N., Pyne C., Babluk L.A., Tikoo S.K.;
 RT "Characterization of early region 1 and p1x of bovine adenovirus-3.";
 RL Virology 253:299-308 (1999).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M8R-1;
 RA Reddy P.S., Idamakanti N., Zakharchouk A.N., Baxi M.K., Lee J.B.,
 RA Pyne C., Babluk L.A., Tikoo S.K.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF030154; AAD09724.1; -;
 DR InterPro; IPR004912; Adeno VII;
 DR Pfam; PF03228; Adeno VII; 1.
 SQ SEQUENCE 171 AA; 16959 MW; 0EACB1C3C12519A CRC64;

Query Match 40.3%; Score 76.5; DB 12; Length 171;
 Best Local Similarity 50.0%; Pred. No. 0.088;
 Matches 21; Conservative 4; Mismatches 10; Indels 7; Gaps 1;

OY 1 RRVVRRVRRV-----VRRVRRVRRVRRVRRVRRV 35
 DB 88 RRGVRRVRLRRSPRLQRRVRRVRRVRRVRRVRRV 129

RESULT 2

O94ML3 PRELIMINARY; PRT; 242 AA.
 AC O94ML3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Membrane protein p6.
 GN 6.
 OS bacteriophage phi-12.
 OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
 OX NCBI_TaxID=161736;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gotschlich P., Wei H., Toporovsky I.;
 RT "Nucleotide sequence of the middle dsRNA segment of Phi-12.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY039807; AAK74123.1; -;
 SQ SEQUENCE 242 AA; 25478 MW; 8F6676077899AF60 CRC64;

Query Match 40.0%; Score 76; DB 9; Length 242;
 Best Local Similarity 30.0%; Pred. No. 0.14;
 Matches 12; Conservative 18; Mismatches 10; Indels 0; Gaps 0;

OY 1 RRVVRRVRRVRRVRRVRRVRRVRRVRRVRRV 40
 DB 5 KSIKAIKSVKAVKAVKAVKAVKAVKAVKAVK 44

RESULT 3

O9SHX2 PRELIMINARY; PRT; 349 AA.
 AC O9SHX2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Atg906420 protein (Hypothetical 41.9 kDa protein).
 GN Atg906420.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RL Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Romling C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhagen G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768 (1999).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanaken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Romling C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nieman W.C.,
 RA Fraser C.M., Venter J.C.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007235; AAD26890.1; -;
 DR EMBL; AC006918; AAM15311.1; -;
 KM Hypothetical protein.
 SQ SEQUENCE 349 AA; 41935 MW; DCE334C856C9F0F5 CRC64;

Query Match 37.6%; Score 71.5; DB 10; Length 349;
 Best Local Similarity 41.9%; Pred. No. 0.62;
 Matches 18; Conservative 10; Mismatches 14; Indels 1; Gaps 1;

OY 1 RRVVRRVRRVRRVRRVRRVRRVRRVRRVRRV 42
 DB 278 RRVVRRVRRVRRVRRVRRVRRVRRVRRVRRV 320

RESULT 4

O9FZT7 PRELIMINARY; PRT; 230 AA.
 AC O9FZT7;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE Membrane protein.
 GN 6.
 OS Pseudomonas bacteriophage phi-13.
 OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
 OX NCBI_TaxID=134554;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Mindich L., Qiao X., Qiao J., Onodera S., Romaneschnk M.,
 RA Hoogerstraaten D.;
 RT "Isolation of additional bacteriophages with genomes of segmented
 double-stranded RNA.";
 RL J. Bacteriol. 181:4505-4508 (1999).
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=20473938; PubMed=11017801;
 RA Qiao X., Qiao J., Onodera S., Mindich L.;
 RT "Characterization of phi13, a bacteriophage related to phi6 and
 RT containing three dsRNA genomic segments.";
 RL Virology 275:218-224 (2000).
 DR EMBL; AF261667; AAG00439.1; -;
 SQ SEQUENCE 230 AA; 24117 MW; 729BAAAF55566256 CRC64;

Query Match 36.8%; Score 70; DB 9; Length 230;
 Best Local Similarity 35.3%; Pred. No. 0.62;
 Matches 12; Conservative 15; Mismatches 7; Indels 0; Gaps 0;

OY 7 RRVVRRVRRVRRVRRVRRVRRVRRVRRVRRV 40
 DB 11 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRRV 44


```
CC -1-SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X71334; CA50474.1; -.
CC PIR; S33332; S33332.
CC InterPro: IPR000492; Protamine_P2.
CC Pfam; PF00841; Protamine_P2; 1.
CC Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
CC Testis; DNA condensation; Nuclear protein.
CC PROPEP 1 ?
CC CHAIN ? 102 SPERM HISTONE P2.
CC SEQUENCE 102 AA; 12976 MW; B10B48BB8CF64C4 CRC64;

Query Match 28.2%; Score 53.5; DB 1; Length 102;
Best Local Similarity 43.9%; Pred. No. 2.2;
Matches 18; Conservative 4; Mismatches 16; Indels 3; Gaps 1;

QY 5 RRVRRVRRVRRVRRVRRV--RRVRRVRRVRRVRR 42
Db 61 RLRLRIHQOHRSCTRRKRKRRSCRRHRRHRRGCTRRRTCRK 101
```

Search completed: June 9, 2003, 11:56:38
Job time : 8.48936 secs

FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 856 AA, 97188 MW, 3373688B84C1AAC CRC64;

Query Match 28.4%; Score 54; DB 1; Length 856;
 Best Local Similarity 35.7%; Pred. No. 16;
 Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 6 RRVRRVRRVRRVRRVRRVRRVRRV 33
 DB 828 RVLEVVQEAIRAIRHPRIRIOGLERIL 855

RESULT 11
 PRT2_CLUPA STANDARD; PRT; 31 AA.
 AC P02336;
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Protamine 2 (Clupeine Z).
 OS Clupea pallasi (Pacific herring), and
 OS Clupea harengus (Atlantic herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
 OC Clupea.
 NCBI_TaxID=30724, 7950;
 RX MEDLINE=7157437; PubMed=5551645;
 RA Iwai K., Nakahara C., Ando T.;
 RT "Studies on protamines. XV. The complete amino acid sequence of the Z
 component of clupeine. Application of N leads to O acyl rearrangement
 and selective hydrolysis in sequence determination.";
 RL J. Biochem. 69:493-509(1971).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=C.pallasi;
 RA Chang W.-J., Nukushina M., Ishii S., Nakahara C., Ando T.;
 RL Submitted (AUG-1970) to the PIR data bank.
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=92126280; PubMed=1772633;

RA Balchmidt P., Hansen F.B., Dodson E., Dodson G., Korber F.;
 RT "Structure of porcine insulin cocrystallized with clupeine Z.";
 RL Acta Crystallogr. B 47:975-986(1991).
 CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- MISCELLANEOUS: CLUPEINE Z IS PROBABLY THE RESULT OF A CROSSOVER
 CC BETWEEN THE GENES FOR CLUPEINES YI AND YII.
 DR PIR; A02678; CHR2.
 DR PDB; 7INS; 31-UN-94.
 KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
 KW Testis; DNA condensation; Nuclear protein; 3D-structure.
 SQ SEQUENCE 31 AA; 4165 MW; 092CCBF7F3AFC050 CRC64;

Query Match 28.2%; Score 53.5; DB 1; Length 31;
 Best Local Similarity 54.3%; Pred. No. 0.67;
 Matches 19; Conservative 0; Mismatches 9; Indels 7; Gaps 2;

QY 8 RRVRRVRRVRRVRRVRRVRRVRRVRR 42
 DB 3 RRSRRASRPVR--RRPRVSR--RRARR 30

RESULT 12
 HSP3_HORSE STANDARD; PRT; 58 AA.
 AC P15313;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sperm histone P2B (STR2B).
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 NCBI_TaxID=9796;
 RX MEDLINE=90304188; PubMed=2364093;
 RA Pirhonen A., Valtanen P., Linnala-Kankkunen A., Heiskanen M.-L.,
 RA Maenpaa P.K.;
 RT "Primary structures of two protamine 2 variants (Str2a and Str2b) from
 RT stallion spermatozoa.";
 RL Biochim. Biophys. Acta 1039:177-180(1990).
 RN [2]
 RP SEQUENCE OF 1-25.
 RX MEDLINE=89171259; PubMed=2924903;
 RA Pirhonen A., Linnala-Kankkunen A., Maenpaa P.K.;
 RT "Comparison of partial amino acid sequences of two protamine 2
 RT variants from stallion sperm. Structural evidence that the variants
 RT are products of different genes.";
 RL FEBS Lett. 244:199-202(1989).
 CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.
 DR PIR; S02787; S02787.
 DR PIR; S10755; S10755.
 KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
 KW Testis; DNA condensation; Nuclear protein.
 SQ SEQUENCE 58 AA; 7979 MW; 8D31F36098A73179 CRC64;

Query Match 28.2%; Score 53.5; DB 1; Length 58;
 Best Local Similarity 53.8%; Pred. No. 1.3;
 Matches 21; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

QY 5 RRVRRVRRVRRVRRVRRVRRVRRVRR 42
 DB 15 RLVYLRRRRRYSSRRRRRPPCR--RRHRYCRRVRR 51


```

SQ      SEQUENCE    102 AA;   12976 MW;   966C7AAB8BCFF4DB CRC64;

Query Match          30.3%; Score 57.5; DB 1; Length 102;
Best Local Similarity 46.3%; Pred. No. 0.75;
Matches 19; Conservative 3; Mismatches 16; Indels 3; Gaps 1;

QY      RRVRRVRRVVRRVVRRV---RVRVRRVRRVVRRVRR 42
        |||::||:|:|||||::|||::|||::|||
Db       RLRLRIHQHRSCKRRKRSCKRRKRKRRCGCTRTTCRR 101

RESULT 5
VC07_ADEB2
ID     VC07_ADEB2      STANDARD; PRT; 183 AA.
AC     Q966Z4;
DT     01-NOV-1997 (Rel. 35, Created)
DT     01-NOV-1997 (Rel. 35, Last sequence update)
DT     16-OCT-2001 (Rel. 40, Last annotation update)
DE     Major core protein precursor (Protein VII) (pVII).
GN     pVII.
OS     Bovine adenovirus type 2 (Mastadenovirus boe2).
OC     Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX     NCBI_TaxID=114429;
RN     [1]
RP     SEQUENCE FROM N.A.
RA     Ruvaai M., Harrach B., Baurevi A., Evans P., Benko M.;
RL     Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC     -----
CC     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC     between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC     the European Bioinformatics Institute. There are no restrictions on its
CC     use by non-profit institutions as long as its content is in no way
CC     modified and this statement is not removed. Usage by and for commercial
CC     entities requires a license agreement (See http://www.isb.ch/announce/
CC     or send an email to licenses@isb-stb.ch).
CC     -----
DR     EMBL; U41423; AB16756.1; -.
DR     InterPro; IPR004912; Adeno_VII.
DR     Pfam; PF03228; Adeno_VII; 1.
KW     Core protein; Late protein.
FT     PROPEP           1         23
FT     CHAIN            24        183
FT     SITE             23         24
FT                                     MAJOR CORE PROTEIN.
FT                                     CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT                                     (POTENTIAL).
SQ      SEQUENCE    183 AA;   20710 MW;   D694365A50D688B CRC64;

Query Match          30.0%; Score 57; DB 1; Length 183;
Best Local Similarity 38.2%; Pred. NO. 1.5;
Matches 21; Conservative 1; Mismatches 19; Indels 14; Gaps 2;

QY      1 RRVRRVRRV-----RVRRVRRVRRVRRVRRVRR 41
        |||::||:|:|||||::|||::|||::|||
Db       RRAERRRRRGATSMRAPRALVLVSARRRLARCGVRRTRTPVADVVAIVEAIR 132

RESULT 6
PRTB_MUGCE
ID     PRTB_MUGCE      STANDARD; PRT; 33 AA.
AC     P08130;
DT     01-AUG-1988 (Rel. 08, Created)
DT     01-AUG-1988 (Rel. 08, Last sequence update)
DT     16-OCT-2001 (Rel. 40, Last annotation update)
DE     Proctamine M6/M7 (Mugilina beta).
OS     Mugil cephalus (Flathead mullet) (Mugil japonicus).
OC     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC     Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC     Acanthomorpha; Acanthopterygii; Percomorpha; Mugilomorpha; Mugilidae;
OC     Mugil.
OX     NCBI_TaxID=48193;
RN     [1]
RP     SEQUENCE.
RA     Tissue=Sperm;
RC     MEDLINE=87279969; PubMed=3301825;
RX
```

```

RA Okamoto Y., Muta E., Ota S.;
RL "Primary structures of M6 and M7 of mugilic beta (Mugil japonicus).",
CC J. Biochem. 101:1017-1024 (1987).
CC -1- FUNCTION: PROTAGONISTS SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SEMM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SEMM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- MISCELLANEOUS: THE SEQUENCE OF COMPONENT M6 IS SHOWN.
DR PIR: A26762; A26762.
DR PIR: A26762; A26762.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
FT Testis; DNA condensation; Nuclear protein.
FT VARIANT 6 6 E -> Q (IN M7 COMPONENT).
FT VARIANT 22 22 I -> M.
SQ SEQUENCE 33 AA; 4473 MW; 4B407DE638A0D29E CRC64;

Query Match 29.5%; Score 56; DB 1; Length 33;
Best Local Similarity 51.9%; Pred. No. 0.37;
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 5 RRVRRVVRRVVRRVVRRVVRR 31
Db 5 RERSPIRRRRARRARRRRRRVR 31
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |

RESULT 7
URK_BACHD STANDARD; PRT; 211 AA.
ID URK_BACHD
O9K0D8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR BH1275.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AP001511; BAB04994.1; -.
DR InterPro; IPR001324; PRK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00478; PHRIBLKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk.1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
NP BIND 12 19 ATP (POTENTIAL).
FT 211 AA; 24387 MW; C2A9A2CB0030520B CRC64;.
SQ SEQUENCE

```

```
OC Alouatta.
OX NCBI_TaxID=9503;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93292525; PubMed=8513810;
RA Retief J.D., Dixon G.H.;
RT "Evolution of pro-protamine P2 genes in primates.";
RL Eur. J. Biochem. 214:609-615(1993).
RN [2]
RP ERRATUM.
RX MEDLINE=94109373; PubMed=8281927;
RA Retief J.D., Dixon G.H.;
RL Eur. J. Biochem. 218:1095-1095(1993).
CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; X71335; CA50475.1; -.
DR PIR; S33338; S33338.
DR InterPro; IPR000492; Protamine P2.
DR Pfam; PF00841; protamine P2; 1.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
FT PROPEP
FT CHAIN
FT PROPEP 1 100 SPERM HISTONE P2.
SQ SEQUENCE 100 AA; 12560 MW; CBF579527B9C6A8 CRC64;
Query Match 31.6%; Score 60; DB 1; Length 100;
Best Local Similarity 50.0%; Pred. No. 0.38;
Matches 19; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
QY 5 RRVRRVRRVRRVRRVRRVRRVRRVRRVRR 42
DB 62 RRLVRRRRQRRCRRRCRRRRNRRCRRTRTCRR 99
RESULT 3
VC07_ADE40 STANDARD; PRT; 185 AA.
ID VC07_ADE40
AC 089533;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
GN pVII.
OS Human adenovirus type 40.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28284;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RX MEDLINE=94087748; PubMed=8263936;
RA Davidson A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
RL "The DNA sequence of adenovirus type 40.";
RT J. Mol. Biol. 234:1308-1316(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RA Pieniazek N.J., Slemenda S.B., Pieniazek D., Luftig R.B.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; L19443; AAC13963.1; -.
DR EMBL; M86655; AAA42526.1; -.
DR InterPro; IPR004912; Adeno VII.
DR Pfam; PF03228; Adeno VII; 1.
FT PROPEP 1 23 BY SIMILARITY.
FT CHAIN 24 185 MAJOR CORE PROTEIN.
FT SITE 23 24 CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT SITE (POTENTIAL).
SQ SEQUENCE 185 AA; 20518 MW; 4FB80E53EF218A9E CRC64;
Query Match 31.6%; Score 60; DB 1; Length 185;
Best Local Similarity 37.5%; Pred. No. 0.69;
Matches 15; Conservative 7; Mismatches 12; Indels 6; Gaps 1;
QY 1 RRVRRVRRVRRVRRVRRVRRVRRVRRVRR 34
DB 92 RRYAQRRLQRRLRRRPTAMTARAVLRRAQRIGRARR 131
RESULT 4
HSP2_PANTR STANDARD; PRT; 102 AA.
ID HSP2_PANTR
AC P35300;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm histone P2 precursor (Protamine P2).
GN PM2.
OS Pan troglodytes (chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93292525; PubMed=8513810;
RA Retief J.D., Dixon G.H.;
RT "Evolution of pro-protamine P2 genes in primates.";
RL Eur. J. Biochem. 214:609-615(1993).
RN [2]
RP ERRATUM.
RX MEDLINE=94109373; PubMed=8281927;
RA Retief J.D., Dixon G.H.;
RL Eur. J. Biochem. 218:1095-1095(1993).
CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; X72968; CA51474.1; -.
DR PIR; S33331; S33331.
DR InterPro; IPR000492; Protamine P2.
DR Pfam; PF00841; protamine P2; 1.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
FT PROPEP 1 102 SPERM HISTONE P2.
FT CHAIN
```


C;Keywords: DNA binding; nucleus

Query Match

Best Local Similarity 29.5%; Score 56; DB 2; Length 33;
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

Qy

5 RRVRRVRRVRRVRRVRRVRRVRR 31

Db

5 RETSRPIRRRRRARAPIRRRRRVRR 31

RESULT 13

C83809

uridine kinase udk [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: C83809

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: C83809

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-211 <STO>

A;Cross-references: GB:AP001511; GB:BA000004; NID:910173727; PIDN:BAB04994.1; GSPDB:GN00

C;Genetics:

A;Gene: udk

C;Superfamily: uridine kinase

Query Match

Best Local Similarity 29.5%; Score 56; DB 2; Length 211;
Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy

9 RRVRRVRRVRRVRRVRRVRRVRRVRR 37

Db

138 RIRRMVDRIRRGRTLESVIEGYTKVRR 166

RESULT 14

CLHRY2

protamine YII - Pacific herring

N;Alternate names: clupeine

C;Species: Clupea pallasii, Clupea harengus pallasii (Pacific herring)

C;Date: 12-Aug-1981 #sequence_revision 12-Aug-1981 #text_change 16-Feb-1997

C;Accession: A38052; A02677

R;Suzuki, K.; Ando, T.

J. Biochem. 72, 1419-1432, 1972

A;Title: Studies on protamines. XVI. The complete amino acid sequence of clupeine YII.

A;Reference number: A38052; MUID:73223106; PMID:4664740

A;Accession: A38052

A;Molecule type: protein

A;Residues: 1-30 <SDZ>

C;Superfamily: protamine Y2

C;Keywords: chromosomal protein; DNA binding; spermatogenesis

Query Match

Best Local Similarity 29.2%; Score 55.5; DB 1; Length 30;
Matches 19; Conservative 0; Mismatches 9; Indels 7; Gaps 2;

Qy

8 RRVRRVRRVRRVRRVRRVRRVRRVRR 42

Db

2 RRRTRASRPVRR--RRPRVSRR-----RRARRR 29

RESULT 15

CLHR2A

protamine YII - Atlantic herring

N;Alternate names: clupeine

C;Species: Clupea harengus, Clupea harengus harengus (Atlantic herring)

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Feb-1997

C;Accession: A37575; A02677

R;Chang, W.J.; Nukushina, M.; Ishii, S.; Nakahara, C.; Ando, T.

submitted to the Atlas, August 1970

A;Reference number: A37575

A;Accession: A37575

A;Molecule type: protein

A;Residues: 1-30 <CHA>

C;Superfamily: protamine Y2

C;Keywords: chromosomal protein; DNA binding; spermatogenesis

Query Match

Best Local Similarity 29.2%; Score 55.5; DB 1; Length 30;
Matches 19; Conservative 0; Mismatches 9; Indels 7; Gaps 2;

Qy

8 RRVRRVRRVRRVRRVRRVRRVRRVRR 42

Db

2 RRRTRASRPVRR--RRPRVSRR-----RRARRR 29

Search completed: June 9, 2003, 12:03:11

Job time : 18.4255 secs

FILING DATE: 11-08-93
 ATTORNEY/AGENT INFORMATION:
 NAME: HULTQUIST, STEVEN J.
 REGISTRATION NUMBER: 28021
 REFERENCE/DOCKET NUMBER: 4013-106
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919)990-9531
 TELEFAX: (919)990-9532
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 37
 TYPE: AMINO ACID
 TOPOLOGY: LINEAR
 MOLECULE TYPE:
 DESCRIPTION: PEPTIDE
 HYPOTHETICAL: NO
 FRAGMENT TYPE: COMPLETE PEPTIDE
 ORIGINAL SOURCE: SYNTHETIC
 IMMEDIATE SOURCE: SYNTHETIC
 PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
 US-08-231-730A-23

Query Match 35.2%; Score 57; DB 1; Length 37;
 Best Local Similarity 22.9%; Pred. No. 0.14; 11; Indels 0; Gaps 0;
 Matches 8; Conservative 16; Mismatches 11; Indels 0; Gaps 0;
 Qy 2 RRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
 Db 3 KKFVKVAKVAKVAKVAKVAKVAKVAKVAKVAK 37

Search completed: June 9, 2003, 12:05:07
 Job time : 13.2553 secs

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-160

Query Match 36.4%; Score 59; DB 1; Length 28;
Best Local Similarity 39.3%; Pred. No. 0.058;
Matches 11; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 7 RVRVVRVRRVRRVRRVRRVRRV 34
DB 1 RVRVVGACRAIRHPRIRIQGLRRL 28

RESULT 8
US-08-932-682-160
Sequence 160, Application US/08932682
Patent No. 5945507
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mietzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,682
FILING DATE: 18-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/786,748
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-705-5000
TELEFAX: 212-765-2519

INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5945507e
US-08-932-682-160

Query Match 36.4%; Score 59; DB 2; Length 28;
Best Local Similarity 39.3%; Pred. No. 0.058;
Matches 11; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 7 RVRVVRVRRVRRVRRVRRVRRV 34
DB 1 RVRVVGACRAIRHPRIRIQGLRRL 28

RESULT 9
US-08-505-486-65
Sequence 65, Application US/08505486
Patent No. 5955573
GENERAL INFORMATION:
APPLICANT: Jesse M. Jaynes
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,486
FILING DATE: 21-JUL-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/279,472
FILING DATE: 22-JUL-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-117A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-505-486-65

Query Match 35.2%; Score 57; DB 2; Length 27;
Best Local Similarity 30.0%; Pred. No. 0.1;
Matches 6; Conservative 12; Mismatches 2; Indels 0; Gaps 0;

COMPUTER READABLE FORM:
COMPUTER TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,174A
FILING DATE: 12-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/079,512
FILING DATE: 18-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bodrowicz, Donna
REGISTRATION NUMBER: 32,196
REFERENCE/DOCKET NUMBER: 0234R2D-US
TELEPHONE: (515) 248-4896
TELEFAX: (515) 334-6883
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-174A-9

Query Match 39.5%; Score 64; DB 1; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.015;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 6 RRVVRRVRRVRRVRRVRRVRRVRRV 33
DB 1 RRVVRRVRRVRRVRRVRRVRRVRRV 28

RESULT 5
PCT-US95-00062-9
Sequence 9, Application PC/TUS9500062
GENERAL INFORMATION:
APPLICANT: Pioneer Hi-Bred International, Inc.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/Microsoft Windows
SOFTWARE: Microsoft Windows Notepad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00062
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Yates, Michael E.; Sweeney, Patricia A.;
Roth, Michael J.; & Simon, Soma G.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 234R2-PCT
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00062-9

Query Match 39.5%; Score 64; DB 5; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.015;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 6 RRVVRRVRRVRRVRRVRRVRRVRRV 33
DB 1 RRVVRRVRRVRRVRRVRRVRRVRRV 28

RESULT 6
US-09-413-814-78
Sequence 78, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloeker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 78
LENGTH: 882
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-78

Query Match 38.0%; Score 61.5; DB 4; Length 882;
Best Local Similarity 53.1%; Pred. No. 0.91;
Matches 17; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 6 RRVVRRVRRVRRVRRVRRVRRVRRV 36
DB 594 RRVVRRVRRVRRVRRVRRVRRVRRV 625

RESULT 7
US-08-786-748A-160
Sequence 160, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Metzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228

PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
STREET: N/A
US-08-436-703B-17

Query Match 40.7%; Score 66; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 0.01;
Matches 15; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6 RRVVVRRVVRRVVRRVVRRVVRRV 35
DB 3 RRAARRARRARRARRARRARRARRR 32

RESULT 2
US-08-436-703B-5
Sequence 5, Application US/08436703B
Patent No. 5919761

GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR
TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
TITLE OF INVENTION: WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 6601 Woodward Avenue
STREET: Suite 1525
CITY: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44mb, 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6;
SOFTWARE: ASCII (DOS)text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7WK-060548-00233
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-5

Query Match 40.7%; Score 66; DB 2; Length 39;
Best Local Similarity 50.0%; Pred. No. 0.01;
Matches 15; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6 RRVVVRRVVRRVVRRVVRRVVRRV 35
DB 3 RRAARRARRARRARRARRARRARRR 32

RESULT 3
US-08-179-632-9
Sequence 9, Application US/08179632
Patent No. 5607914

GENERAL INFORMATION:
APPLICANT: Rao, A. Gururaj, Zhong, Lingxiu
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/Microsoft Windows
SOFTWARE: Microsoft Windows No. 5607914epad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/179,632
FILING DATE: 07-JAN-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/079,512
FILING DATE: 06/18/93
ATTORNEY/AGENT INFORMATION:
NAME: Roth, Michael J.
REGISTRATION NUMBER: 29,342
REFERENCE/DOCKET NUMBER: 0233 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 245-3594
TELEFAX: (515) 245-3634
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-179-632-9

Query Match 39.5%; Score 64; DB 1; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.015;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 6 RRVVVRRVVRRVVRRVVRRVVRRV 33
DB 1 RRIYRAIRHPIPRIRIGMLRRIGRIERV 28

RESULT 4
US-08-440-174A-9
Sequence 9, Application US/08440174A
Patent No. 5717061
GENERAL INFORMATION:
APPLICANT: Rao, Gururaj A.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: 7100 N.W. 62nd Avenue
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:50:31 ; Search time 12.2553 Seconds
(without alignments)
86.430 Million cell updates/sec

Title: US-10-079-075-6

Perfect score: 162
Sequence: 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	40.7	38	2	US-08-436-703B-17
2	66	40.7	39	2	US-08-436-703B-5
3	64	39.5	31	1	US-08-179-632-9
4	64	39.5	31	1	US-08-440-174A-9
5	64	39.5	31	5	PCT-US95-00062-9
6	61.5	38.0	882	4	US-09-413-814-78
7	59	36.4	28	1	US-08-786-748A-160
8	59	36.4	28	2	US-08-932-682-160
9	57	35.2	27	2	US-08-505-486-65
10	57	35.2	27	3	US-08-801-028-65
11	57	35.2	27	3	US-09-340-154-65
12	57	35.2	27	4	US-09-482-611B-65
13	57	35.2	27	5	PCT-US95-09339-65
14	57	35.2	27	5	PCT-US95-09339-65
15	57	35.2	27	1	US-08-231-730A-23
16	57	35.2	37	1	US-08-427-001C-23
17	57	35.2	37	1	US-08-457-798-23
18	57	35.2	37	1	US-08-457-171-23
19	57	35.2	37	2	US-08-505-486-23
20	57	35.2	37	3	US-08-689-489C-23
21	57	35.2	37	3	US-08-801-028-23
22	57	35.2	37	3	US-09-340-154-23
23	57	35.2	37	4	US-09-232-802A-23
24	57	35.2	37	4	US-09-482-611B-23
25	57	35.2	37	5	PCT-US94-06176-23
26	57	35.2	37	5	PCT-US94-12550-23
27	57	35.2	37	5	PCT-US95-04335-23

28	57	35.2	37	5	PCT-US95-04718-23	Sequence 23, Appl
29	57	35.2	37	5	PCT-US95-09338-23	Sequence 23, Appl
30	57	35.2	37	5	PCT-US95-09338-23	Sequence 23, Appl
31	56.5	34.0	105	2	US-08-668-255-5	Sequence 5, Appl
32	55	34.0	27	2	US-08-505-486-64	Sequence 64, Appl
33	55	34.0	27	3	US-08-801-028-64	Sequence 64, Appl
34	55	34.0	27	3	US-09-340-154-64	Sequence 64, Appl
35	55	34.0	27	4	US-09-482-611B-64	Sequence 64, Appl
36	55	34.0	27	5	PCT-US95-09339-64	Sequence 64, Appl
37	55	34.0	27	5	PCT-US95-09339-64	Sequence 64, Appl
38	55	34.0	28	1	US-08-786-748A-18	Sequence 18, Appl
39	55	34.0	28	2	US-08-932-682-18	Sequence 18, Appl
40	54.5	33.6	96	2	US-08-668-255-7	Sequence 7, Appl
41	54	33.3	28	1	US-08-786-748A-14	Sequence 14, Appl
42	54	33.3	28	1	US-08-786-748A-19	Sequence 19, Appl
43	54	33.3	28	1	US-08-786-748A-24	Sequence 24, Appl
44	54	33.3	28	1	US-08-786-748A-27	Sequence 27, Appl
45	54	33.3	28	1	US-08-786-748A-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-08-436-703B-17
Sequence 17, Application US/08436703B
Patent No. 5919761
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR
TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
TITLE OF INVENTION: WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benita J, Rohm, Esq.
STREET: 6601 Woodward Avenue
STREET: Suite 1525
CITY: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6/
SOFTWARE: ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7WK-060548-00233
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A

```
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A 34001 / 072396.0222
/ CURRENT APPLICATION NUMBER: US/09/785,058
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-11

Query Match      82.7%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 5.6e-11;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 VRRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB      1 VRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 13
US-09-785-059-12
/ Sequence 12, Application US/09785059
/ Patent No. US20020169279A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A3577 / 072396.0217
/ CURRENT APPLICATION NUMBER: US/09/785,059
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 12
/ LENGTH: 48
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-059-12

Query Match      75.9%; Score 123; DB 9; Length 48;
Best Local Similarity 90.6%; Pred. No. 1.9e-09;
Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VRRVRRVRRVRRVRRVRRVRRVRRVRR 32
DB      13 VRRVRRVRRVRRVRRVRRVRRVRRVRR 44

RESULT 14
US-10-079-075-12
/ Sequence 12, Application US/10079075
/ Publication No. US20020188102A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A34001-A / 072396.0222
/ CURRENT APPLICATION NUMBER: US/10/079,075
/ CURRENT FILING DATE: 2002-02-19
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 12
/ LENGTH: 48
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
```

```
/ FEATURE:
/ OTHER INFORMATION: artificial peptides derived from HIV-1
US-10-079-075-12

Query Match      75.9%; Score 123; DB 9; Length 48;
Best Local Similarity 90.6%; Pred. No. 1.9e-09;
Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VRRVRRVRRVRRVRRVRRVRRVRRVRR 32
DB      13 VRRVRRVRRVRRVRRVRRVRRVRRVRR 44

RESULT 15
US-09-785-058-12
/ Sequence 12, Application US/09785058
/ Publication No. US20030036627A1
/ GENERAL INFORMATION:
/ APPLICANT: Ronald C. Montelaro
/ APPLICANT: Timothy A. Metzner
/ TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
/ FILE REFERENCE: A 34001 / 072396.0222
/ CURRENT APPLICATION NUMBER: US/09/785,058
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 12
/ LENGTH: 48
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: artificial peptides derived from HIV-1
US-09-785-058-12

Query Match      75.9%; Score 123; DB 9; Length 48;
Best Local Similarity 90.6%; Pred. No. 1.9e-09;
Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VRRVRRVRRVRRVRRVRRVRRVRRVRR 32
DB      13 VRRVRRVRRVRRVRRVRRVRRVRRVRR 44

Search completed: June 9, 2003, 12:34:09
Job time : 19.1489 secs
```

FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785,059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-8

Query Match 100.0%; Score 162; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 13 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 48

RESULT 8
US-10-079-075-8

Sequence 8, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079,075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-8

Query Match 100.0%; Score 162; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 13 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 48

RESULT 9

US-09-785-058-8
Sequence 8, Application US/09785058
Publication No. US20030036627A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A 34001 / 072396.0222
CURRENT APPLICATION NUMBER: US/09/785,058
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-8

Query Match 100.0%; Score 162; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 13 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 48

RESULT 10

US-09-785-059-11
Sequence 11, Application US/09785059
Patent No. US20020169279A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785,059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-11

Query Match 82.7%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 5.6e-11;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 11

US-10-079-075-11
Sequence 11, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Mietzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079,075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-11

Query Match 82.7%; Score 134; DB 9; Length 36;
Best Local Similarity 88.9%; Pred. No. 5.6e-11;
Matches 32; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 12

US-09-785-058-11
Sequence 11, Application US/09785058
Publication No. US20030036627A1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 12:01:36 ; Search time 19.1489 Seconds
(without alignments)
194.092 Million cell updates/sec

Title: US-10-079-075-6
Perfect score: 162
Sequence: 1 VRRVRRVRRVRRVRRVRRVRRVRRVRR 36

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	162	100.0	36	US-09-785-059-6	Sequence 6, App11
2	162	100.0	36	US-10-079-075-6	Sequence 6, App11
3	162	100.0	36	US-09-785-058-6	Sequence 6, App11
4	162	100.0	42	US-09-785-059-7	Sequence 7, App11
5	162	100.0	42	US-10-079-075-7	Sequence 7, App11
6	162	100.0	42	US-09-785-058-7	Sequence 7, App11
7	162	100.0	48	US-09-785-059-8	Sequence 8, App11
8	162	100.0	48	US-10-079-075-8	Sequence 8, App11
9	162	100.0	48	US-09-785-058-8	Sequence 8, App11
10	162	100.0	36	US-09-785-059-11	Sequence 11, App1
11	134	82.7	36	US-10-079-075-11	Sequence 11, App1
12	134	82.7	36	US-09-785-058-11	Sequence 11, App1
13	123	75.9	48	US-09-785-059-12	Sequence 12, App1
14	123	75.9	48	US-10-079-075-12	Sequence 12, App1
15	123	75.9	48	US-09-785-058-12	Sequence 12, App1
16	109	67.3	24	US-09-785-059-5	Sequence 5, App11
17	109	67.3	24	US-10-079-075-5	Sequence 5, App11
18	109	67.3	24	US-09-785-058-5	Sequence 5, App11
19	88	54.3	9	US-09-785-059-10	Sequence 10, App1

20	88	54.3	24	9	US-10-079-075-10	Sequence 10, App1
21	88	54.3	24	9	US-09-785-058-10	Sequence 10, App1
22	83	51.2	31	9	US-09-785-059-2	Sequence 2, App11
23	83	51.2	31	9	US-10-079-075-2	Sequence 2, App11
24	83	51.2	31	9	US-09-785-058-2	Sequence 2, App11
25	72	44.4	28	9	US-09-785-059-1	Sequence 1, App11
26	72	44.4	28	9	US-10-079-075-1	Sequence 1, App11
27	72	44.4	28	9	US-09-785-058-1	Sequence 1, App11
28	66	40.7	31	9	US-09-785-059-3	Sequence 3, App11
29	66	40.7	31	9	US-10-079-075-3	Sequence 3, App11
30	66	40.7	31	9	US-09-785-058-3	Sequence 3, App11
31	56.5	34.9	105	9	US-10-093-892-5	Sequence 5, App11
32	54.5	33.6	96	9	US-10-093-892-7	Sequence 7, App11
33	54	33.3	12	9	US-09-785-059-4	Sequence 4, App11
34	54	33.3	12	9	US-10-079-075-4	Sequence 4, App11
35	54	33.3	12	9	US-09-785-058-4	Sequence 4, App11
36	53	32.7	96	9	US-10-093-892-9	Sequence 9, App11
37	51	31.5	42	10	US-09-810-310-5	Sequence 5, App11
38	51	31.5	48	10	US-09-810-310-4	Sequence 4, App11
39	51	31.5	297	9	US-09-738-626-6303	Sequence 6303, App11
40	50	30.9	29	9	US-10-060-102-8	Sequence 8, App11
41	49	30.2	18	10	US-09-840-009-36	Sequence 36, App11
42	49	30.2	18	10	US-09-840-009-37	Sequence 37, App11
43	49	30.2	344	9	US-10-040-349B-1	Sequence 1, App11
44	49	30.2	345	9	US-10-026-741-49	Sequence 49, App11
45	49	30.2	345	10	US-09-779-451-8	Sequence 8, App11

ALIGNMENTS

RESULT 1
US-09-785-059-6
; Sequence 6, Application US/09785059.
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-6

Query Match 100.0%; Score 162; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.7e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

Qy 1 VRRVRRVRRVRRVRRVRRVRRVRRVRR 36
Db 1 VRRVRRVRRVRRVRRVRRVRRVRRVRR 36

RESULT 2
US-10-079-075-6
; Sequence 6, Application US/10079075
; Patent No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mletzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12

OS Synthetic.
 XX W09527497-A1.
 PN
 XX 19-OCT-1995.
 PD
 XX
 PF 06-APR-1995; 95WO-US04335.
 XX
 PR 08-APR-1994; 94US-0225476.
 XX
 XX (DEME-) DEMETER BIOTECHNOLOGIES LTD.
 PA
 PI Jaynes JM, Julian GR;
 DR WPI; 1995-366226/47.
 XX
 XX Method of combatting mammalian neoplasias and other disease states -
 PT by delivering non-naturally occurring, non-onco-cytologically
 PT proliferative lytic peptide
 XX
 XX Claim 7; Page 47; 63pp; English.
 PS
 CC AAR77042-R77081 are synthetic, amphipathic, lytic peptide analogues of
 CC melittin, cecropin, magainin and defensin peptides. The peptides are
 CC between 23 and 39 residues long, are amphipathic, carry an overall
 CC positive charge and have anti-neoplastic activity. The peptides are
 CC specifically useful for the lysis of cancer cells. Normal mammalian
 CC cells are resistant to lysis due to their highly organised
 CC cytoskeleton, cancerous cells however possess an inferior and
 CC structurally compromised cytoskeleton which when acted upon by lytic
 CC peptides will cause cell lysis. This allows the lytic peptides to be
 CC used for in vivo treatment of cancers. The peptides are esp. useful
 CC for the treatment of female mammalian cancers e.g. breast, ovarian,
 CC uterine and cervical cancers. The peptides can however be used to
 CC treat most forms of cancer, cystic fibrosis, pneumonia, bronchitis,
 CC and bronchopulmonary viral and microbial infections.
 XX
 SQ Sequence 37 AA;
 Query Match 35.2%; Score 57; DB 16; Length 37;
 Best Local Similarity 22.9%; Pred. No. 0.94;
 Matches 8; Conservative 16; Mismatches 11; Indels 0; Gaps 0;
 QY 2 RRVRVRVRVRVRVRVRVRVRVRVRVR 36
 DB 3 KKFVKVAKVAKVAKVAKVAKVAKVAKK 37
 RESULT 14
 AAR74713
 ID AAR74713 standard; peptide; 37 AA.
 XX
 AC AAR74713;
 XX
 DT 11-JUN-1996 (first entry)
 XX
 DE Tryptic digestion resistant lytic peptide.
 XX
 XX Tryptic digestion resistant; immunological; infection;
 KM lytic; methylated; tryptic resistant; immunological; infection;
 XX neoplasia.
 OS Synthetic.
 XX
 OS Key Location/Qualifiers
 FT Modified-site 1..37
 FT /note= "the epsilon-amino groups of the lysine
 FT residues and the alpha-amino group of the
 FT N-terminal amino acid are sufficiently
 FT methylated to impart enhanced proteolytic
 FT digestion resistance to the peptide"
 XX
 PN W09513085-A1.
 XX

PD 18-MAY-1995.
 XX
 XX 01-NOV-1994; 94WO-US12550.
 PF
 XX
 PR 08-NOV-1993; 93US-0148889.
 XX
 XX (DEME-) DEMETER BIOTECHNOLOGIES LTD.
 PA
 PI Julian GR;
 DR WPI; 1995-193900/25.
 XX
 XX New lytic peptide(s) for treating e.g. infections and neoplasias -
 PT contain mainly alanine, valine and lysine residues with the lysine
 PT residues pref. methylated
 XX
 XX Claim 8; Page 49; 66pp; English.
 PS
 CC AAR74699-R74733 are synthetic lytic peptides which are rich in the
 CC amino acids lysine, valine and alanine. They are useful in the
 CC treatment of neoplasia and viral, bacterial, protozoan, fungal or
 CC yeast infections. The lysine residues and the N-terminal amino acid
 CC residue of the peptides are methylated which imparts enhanced
 CC resistance to proteolytic digestion and allows effective in vivo
 CC administration.
 XX
 SQ Sequence 37-AA;
 Query Match 35.2%; Score 57; DB 16; Length 37;
 Best Local Similarity 22.9%; Pred. No. 0.94;
 Matches 8; Conservative 16; Mismatches 11; Indels 0; Gaps 0;
 QY 2 RRVRVRVRVRVRVRVRVRVRVRVRVR 36
 DB 3 KKFVKVAKVAKVAKVAKVAKVAKVAKK 37
 RESULT 15
 AAR64792
 ID AAR64792 standard; peptide; 37 AA.
 XX
 AC AAR64792;
 XX
 DT 24-AUG-1995 (first entry)
 XX
 DE Amphipathic peptide #23, for treating a pulmonary disease state.
 XX
 XX Amphipathic peptide; pulmonary disease; resistant; bacteria;
 KM proteolytic digestion; methylation; glyoxylation; cystic fibrosis;
 KM CF; neoplasia; pneumonia; bronchitis; lytic activity; lysis.
 XX
 OS Synthetic.
 XX
 OS Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Opt. methylated or glyoxylated"
 FT Modified-site 1..37
 FT /note= "Lys residues are opt. methylated and/or Arg
 FT residues are glyoxylated"
 XX
 PN W09428921-A.
 XX
 PD 22-DEC-1994.
 XX
 PF 02-JUN-1994; 94WO-US06176.
 XX
 PR 04-JUN-1993; 93US-0039620.
 XX
 PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
 PI Jaynes JM, Julian GR;
 DR WPI; 1995-036106/05.
 XX

XX Disclosure; Column 21; 62pp; English.

XX This sequence represents an antimicrobial peptide of the invention, and
 CC is an analogue of the peptide LRP1 (see AA32549). The peptides can be
 CC used for treating infections caused by *Staphylococcus aureus*,
 CC methicillin resistant *S. aureus*, *Pseudomonas aeruginosa*, *Enterococcus*
 CC faecalis, *S. marcescens*, *Escherichia coli*, fungi, protozoa and viruses in
 CC a mammalian host. They can be used to inhibit growth of diverse
 CC microorganisms such as bacteria, fungi, protozoa and DNA and RNA viruses
 CC and can be used in tissue culture to inhibit unwanted microbial growth,
 CC particularly for the production of recombinant proteins or vectors for
 CC gene therapy. They can also be used in preventing infections through the
 CC sterilization of wounds prior to suture and to sterilise surgical
 CC instruments. The unique structure of these antimicrobial peptides
 CC imparts high potency while selectivity is maintained, they are
 CC moderately haemolytic but only lyse red blood cells at high
 CC concentrations unlike melittin, a peptide extracted from bee venom, which
 CC is highly active against bacteria and lyses red blood cells showing
 CC little selectivity. The peptides target a membrane structure which makes
 CC it more difficult for a microorganism to develop a mechanism of
 CC resistance against this type of antibiotic. Their small size makes them
 CC relatively simple to prepare by standard synthetic peptide chemistry.

SO Sequence 28 AA;

Query Match 36.4%; Score 59; DB 20; Length 28;
 Best Local Similarity 39.3%; Pred. No. 0.41;
 Matches 11; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 7 RVRVRVRVRVRVRVRVRVRVRV 34
 1 RVRVRVRVRVRVRVRVRVRVRV 28

Db 1 RVRVRVRVRVRVRVRVRVRVRV 28

RESULT 9
 AAW05116
 ID AAW05116 standard; peptide; 17 AA.

XX AAW05116;
 AC
 XX
 DT 19-JUN-1997 (first entry)
 DE Porcine somatotropin mimic.
 XX
 KM Porcine somatotropin; pST; growth promoter; helical conformation.
 XX
 OS Synthetic.
 XX
 PN WO9630405-A1.
 PD 03-OCT-1996.
 XX
 PF 15-MAR-1996; 96WO-US03490.
 PR 31-MAR-1995; 95US-0415239.
 PA (AMCY) AMERICAN CYANAMID CO.
 XX
 PI Buckwalter BL, Shieh H, Wang BS;
 DR WPI; 1996-485447/48.
 PT Peptide(s) mimicking a helical region of porcine somatotropin - used
 XX in compositions to promote mammalian growth
 PS Claim 3; Page 17; 63pp; English.
 CC New peptides are disclosed which, by virtue of having certain
 CC defined amino acids at every third or fourth residue, have a well
 CC defined secondary structure which mimics the helical conformation
 CC of a corresponding region of porcine somatotropin (pST). The peptides
 CC enhance the activity of pST and promote the growth of warm-blooded

CC animals, especially pigs. They compete with pST for binding to the
 CC PS-7.6 monoclonal antibody. The peptides have the generic sequence
 CC XXXIXXXXXXXXV (I) or XXXIXXXXXXXXV (II); where residues X are
 CC undefined other than the statement that the sequences differ from the
 CC native sequence of pST. Formula (II) represents a peptide in which the
 CC location of the essential amino acids is shifted by three amino acids,
 CC representing almost one turn along the helix. Preferably X(2) of (II) is
 CC Ile. Preferably the peptides contain Ser (as a promoter of helical
 CC conformation) as the amino acid immediately amino-terminal to the first
 CC Leu in (I) or to the first Ile of (II). Also, one or more of the first
 CC or second Leu or the Val of (I) may be replaced by Nle. Furthermore, a
 CC Cys residue may be added to either or both ends of the peptides.
 CC The present sequence represents a specific example of the new
 CC peptides.

SO Sequence 17 AA;

Query Match 35.8%; Score 58; DB 17; Length 17;
 Best Local Similarity 60.0%; Pred. No. 0.33;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 12 VRRVRVRVRVRVRV 26
 3 LRRIRRRVRVRIRI 17

Db 3 LRRIRRRVRVRIRI 17

RESULT 10
 AAR92436
 ID AAR92436 standard; peptide; 27 AA.

XX AAR92436;
 AC
 XX
 DT 18-SEP-1996 (first entry)
 DE Lytic peptide used in ubiquitin-lytic peptide fusion protein.
 XX
 KM Ubiquitin; fusion protein; lysis; infection; neoplasia; wound healing;
 XX stability; reduced toxicity.
 XX
 OS Synthetic.
 XX
 PN WO9603519-A1.
 PD 08-FEB-1996.
 XX
 PF 24-JUL-1995; 95WO-US09339.
 PR 22-JUL-1994; 94US-0279472.
 PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
 XX (USDA) US SEC OF AGRIC.
 XX
 PI Belknap W, Garbarino J, Jaynes J;
 DR WPI; 1996-117061/12.
 PT New fusion protein of ubiquitin and a lytic peptide - for treating
 XX infections and neoplasia, healing wounds, etc. also related nucleic
 PT acid, vectors, and transformed cells
 XX
 PS Claim 5; Page 26; 112pp; English.
 CC AAR92372-R92462 are lytic peptides used to create ubiquitin-lytic
 CC peptide fusion proteins in which the ubiquitin polypeptide is linked
 CC at its 3'-terminus to the lytic peptide. The lytic peptides are pref.
 CC selected from either the cecropins, defensins, sarcotoxins, melittin
 CC and magalins. The fusion proteins (FPS) are useful for treating
 CC protozoal, bacterial, fungal and viral infections and neoplasia (in
 CC plants and animals) in the same way as the FP alone, they also
 CC promote wound healing. FPS produced in bacteria may be cleaved in
 CC vitro by ubiquitin hydrolases to recover the active lytic peptide.
 CC FPS produced in eukaryotic cells are cleaved by endogenous enzymes
 CC to yield lytic peptide. Recombinant DNA encoding the FPS have

ID AAP91336 standard; peptide; 28 AA.
XX
AC AAP91336;
XX
DT 19-MAR-1990 (first entry)
XX
DE Amino acid sequence of Shiva-4.
XX
KM Shiva-4; lytic peptide; antimicrobial peptide; disease-resistant
KM trichophyte; Shiva-2; Shiva-3; Shiva-5; Shiva-6; Shiva-7.
XX
XX MO8904371-A.
XX PN
XX 18-MAY-1989.
XX PD
XX 02-NOV-1988; 88MO-US03908.
XX PF
XX 02-NOV-1987; 87US-0115941.
XX PR
XX (LOU) LOUISIANA STATE UNIV.
XX PA
XX Jaynes JM, Derrick KS;
XX PI
XX WPI; 1989-165650/22.
XX DR
XX Transformed plants contg. heterologous gene - expressing antimicrobial
CC agent, or polypeptide high in essential amino acids
CC
XX Table I; ; 56pp; English.
XX PS
XX Amino acid sequence of Shiva-4 as an exemplary lytic peptide for
CC use as an antimicrobial peptide contemplated for use in plant
CC (trichophyte) transformants in the invention. It is a homologue of
CC Shiva-2, -3 and -5 to -7. All of these Shiva peptides are also
CC contemplated as having general utility in inducing lysis of cells in
CC vitro. Shiva-4 may be too lytically active to be used in plants at high
CC expression levels.
CC
XX
SQ Sequence 28 AA;
Query Match 38.9%; Score 63; DB 10; Length 28;
Best Local Similarity 38.7%; Pred. No. 0.13;
Matches 12; Conservative 15; Mismatches 0; Indels 4; Gaps 1;
QY 5 VRRVRRVRRVRRVRRVRRVRRVRRV 35
DB 1 LRRLLRLRLRLRLRLRLRLRLRLRLRL 27
RESULT 7
AAW47769 standard; peptide; 28 AA.
ID AAW47769
XX
XX AAW47769;
XX AC
XX 26-MAY-1998 (first entry)
XX DT
XX Antimicrobial peptide LRP1 analogue.
XX DE
XX Antimicrobial; transmembrane protein; TM; lentivirus lytic peptide;
KM LRP; amphipathic; antibacterial; antifungal; antiviral; antiprotocoa.
XX
XX Synthetic.
OS Human immunodeficiency virus.
OS
XX US5714577-A.
XX PN
XX 03-FEB-1998.
XX PD
XX 24-JAN-1997; 97US-0786748.
XX PF
XX 26-JAN-1996; 96US-0010634.
XX PR
XX 24-JAN-1997; 97US-0786748.
XX PT

XX
XX (UYPI-) UNIV PITTSBURGH.
XX PA
XX Metzner TA, Montelaro RC, Tencza SB;
XX PI
XX WPI; 1998-158352/14.
XX DR
XX Retroviral TM peptides - useful as antibacterial agents
XX PT
XX Disclosure; Column 19; 59pp; English.
XX PS
XX The invention relates to new antimicrobial peptides which correspond to
CC amino acid sequences in the transmembrane proteins of lentiviruses, in
CC particular HIV and SIV. These peptides comprise arginine rich sequences
CC which, when modelled for secondary structure, display high
CC amphipathicity and hydrophobic moment. Also disclosed are structural
CC and functional analogues and homologues of these peptides which also
CC display antimicrobial activity. The peptides are highly inhibitory to
CC microorganisms (bacteria, fungi, viruses and protozoa) but significantly
CC less toxic to red blood cells and other normal mammalian cells. Activity
CC is demonstrated against Gram positive and negative bacteria including
CC Pseudomonas aeruginosa, Staphylococcus aureus, Enterococcus faecalis and
CC Serratia marcescens.
CC
CC The present sequence is one of 169 disclosed specific examples of
CC the new peptides. It is an analogue of the peptide designated LRP1
CC (see AAW47614) which is a peptide from the transmembrane protein (GP41)
CC of HIV strain HXB2R.
XX
SQ Sequence 28 AA;
Query Match 36.4%; Score 59; DB 19; Length 28;
Best Local Similarity 39.3%; Pred. No. 0.41;
Matches 11; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
QY 7 RVRVRRVRRVRRVRRVRRVRRVRRV 34
DB 1 RVRVQACRAIRHPRIRIRIQGLRLRL 28
RESULT 8
AAV32703 standard; peptide; 28 AA.
ID AAV32703
XX
XX AAV32703;
XX AC
XX 21-OCT-1999 (first entry)
XX DT
XX Antimicrobial peptide LRP1 analogue.
XX DE
XX Antimicrobial peptide; LRP1; SLP-1; LRP2; SLP2A; SLP2B; EBP; infection;
KM growth inhibitor; microorganism; virus; gene therapy; vector production;
KM sterilisation.
XX
XX Synthetic.
OS Human immunodeficiency virus type 1.
OS
XX US5945507-A.
XX PN
XX 31-AUG-1999.
XX PD
XX 18-SEP-1997; 97US-0932682.
XX PF
XX 26-JAN-1996; 96US-0010634.
XX PR
XX 24-JAN-1997; 97US-0786748.
XX PR
XX 18-SEP-1997; 97US-0932682.
XX XX
XX (UYPI-) UNIV PITTSBURGH.
XX PA
XX Metzner TA, Montelaro RC, Tencza SB;
XX PI
XX WPI; 1999-508189/42.
XX DR
XX Antimicrobial peptides useful for treating microbial infections
XX PT

reverse heparin anticoagulation. One of the major components of salmine protamine is a 32 amino acid peptide having a total cationic charge of [+21], with arginine accounting for 67% of the total sequence and for all of the positive charge. Peptides of 20-40 amino acids with total cationic charge less than [+21] and which are able, at least partially, to reverse the effect of heparin and/or low molecular weight heparin anticoagulants are claimed. Specifically, the peptides are polycationic analogues of n-protamine where the positive charge on the amino acid sequence is reduced by selective replacement of positively charged arginine residues with an uncharged residue, so that total cationic charge is less than [+21]. The new peptides are used in vivo to reverse the effects of heparin; they have the same anti-heparin activity as protamine but are less toxic (because of the reduced number of positive charges) and are relatively easy and inexpensive to prepare. The present sequence represents a specifically claimed protamine-like peptide with a charge of [+18]; this peptide also includes an RGD fibronectin receptor ligand motif. Peptide [+18RGD] improved the reversal of anti-Factor Xa activity of the low molecular weight heparin Enoxaparin to 72%, compared to only 30% reversal by protamine. Also, peptide [+18RGD] produced less decrease in platelet count than n-protamine.

SQ Sequence 39 AA;

Query Match 40.7%; Score 66; DB 18; Length 39;
Best Local Similarity 50.0%; Pred. No. 0.078;
Matches 15; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6 RRVVRRVRRVRRVRRVRRVRRVRRV 35
DB 3 RRAARRARRARRARRARRARRARRR 32

RESULT 4
AAR60065
ID AAR60065 standard; peptide; 31 AA.

AC AAR60065;

DT 16-FEB-1995 (first entry)

XX 16-FEB-1995 (first entry)

DE Antimicrobial peptide.

XX Amphipathic peptide; alpha-helix; lytic peptide; antifungal;

KM antimicrobial; fungus resistance; disease resistance;

XX crop protection.

OS Synthetic.

XX WO9415961-A.

PN 21-JUL-1994.

PD 12-JAN-1994; 94WO-US00383.

XX 13-JAN-1993; 93US-0003884.

PR (PION-) PIONEER HI-BRED INT INC.

PA Rao AG, Zhong L;

XX WPI; 1994-249137/30.

DR WPI; 1994-249137/30.

PT Synthetic polypeptide(s) and the nucleic acid encoding them -

XX exhibits amphipathic alpha-helices and provide cell-expressable

PT antimicrobial activity

PS Disclosure; Page 19; 25pp; English.

XX The synthetic amphipathic alpha-helical lytic peptides given in

CC AAR60057-71 were designed to provide antifungal or antimicrobial

CC activity when expressed in monocot or dicot plants. They also have

CC veterinary and medical applications.

XX Sequence 31 AA;

Query Match 39.5%; Score 64; DB 15; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.11;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 6 RRVVRRVRRVRRVRRVRRVRRVRRV 33
DB 1 RRIYRAIRHPRIRIGMLRIGRIERY 28

RESULT 5
AAR80735
ID AAR80735 standard; peptide; 31 AA.

AC AAR80735;

DT 28-FEB-1996 (first entry)

XX Synthetic antimicrobial/antifungal polypeptide.

KM Antimicrobial; antifungal; pathogen; plant; amphipathic;

XX broad spectrum.

OS Synthetic.

PN WO9518855-A2.

PD 13-JUL-1995.

XX 06-JAN-1995; 95WO-US00062.

XX 07-JAN-1994; 94US-0179632.

XX (PION-) PIONEER HI-BRED INT INC.

PA Rao AG, Zhong L;

XX WPI; 1995-255059/33.

PT New antimicrobial amphipathic polypeptide(s) and related nucleic
PT acids - for clinical use or esp. to increase resistance of plants
to fungal pathogens.

PS Claim 1; Page 17; 24pp; English.

CC AAR80727-R80748 are new amphipathic polypeptides. They have a broad
CC spectrum of antimicrobial and antifungal activity. They can be used
CC to treat or prevent infection in humans and animals or applied to
CC plants as sprays, creams, dust, etc. The DNA encoding these
CC peptides can also be incorporated into susceptible plants via the
CC use of a non-phytotoxic vehicle adapted for systemic administration.
CC This process imparts resistance to plant pathogens esp. fungi (e.g.
CC Fusarium graminearum, F. moniliforme, Aspergillus flavus, Alternaria
CC longipes, Colletotrichum graminicola, Phytophthora megasperme,
CC sclerotinia sclerotium). The peptides are esp. useful in
CC transformed plants such as maize, sorghum, wheat, soya, alfalfa,
CC rapeseed, sunflower, tobacco or tomato.

SQ Sequence 31 AA;

Query Match 39.5%; Score 64; DB 16; Length 31;
Best Local Similarity 42.9%; Pred. No. 0.11;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 6 RRVVRRVRRVRRVRRVRRVRRVRRV 33
DB 1 RRIYRAIRHPRIRIGMLRIGRIERY 28

RESULT 6
AAP91336

DR WPI; 1995-276981/37.
XX Complex of nucleic acid and oligopeptide with sec. structure - and
PT transfer vectors contg. them, useful for efficient transfer of
PT nucleic acid to cells in gene therapy.
XX
PS Claim 6; Page 16; 20pp; French.
XX
CC The present peptide corresponds to a generic formula for a cationic
CC oligopeptide; the formula is (b-1-l-b)ⁿ, where b is a hydrophobic amino
CC acid, l is a hydrophilic amino acid and n is at least 4. In this case,
CC where b is leu, l is Arg and n = 10, the oligopeptide forms an alpha-
CC helix which forms a stable complex with a nucleic acid. The complex
CC is suitable for transferring nucleic acid, esp. in gene therapy.
XX
SQ Sequence 40 AA;
Query Match 47.2%; Score 76.5; DB 16; Length 40;
Best Local Similarity 46.2%; Pred. No. 0.0041;
Matches 18; Conservative 18; Mismatches 0; Indels 3; Gaps 3;
QY 1 RRVRRRVV-RVRRVRR-RVRRVRRV-RVRRVRR 36
Db 1 LRRLLRRLRLRLRLRLRLRLRLRLRLRLRLRLRL 39
RESULT 2
AAG65539
ID AAG65539 standard; peptide; 18 AA.
XX
AC AAG65539;
XX
DT 30-NOV-2001 (first entry)
XX
DB Peptide sequence used in the course of the invention.
XX
KM Antimicrobial; microbial membrane disrupter; gene therapy; pathogen;
KM spermicide; imaging; magainin; PG1A.
XX
OS Synthetic.
XX
PN WO200160162-A2.
XX
PD 23-AUG-2001.
XX
PF 15-FEB-2001; 2001WO-US04822.
XX
PR 15-FEB-2000; 2000US-0182495.
XX
PA (UYOH-) UNIV OHIO.
XX
PI Blazysk JF;
XX
DR WPI; 2001-565322/63.
XX
PT Novel peptides having antimicrobial activity have positive charge to
PT selectively disrupt microbial membranes, assume beta sheet structure in
PT membrane environment and are substantially amphipathic in beta sheet
PT structure -
XX
PS Example 102; Page 84; 119pp; English.
XX
CC The invention provides an antimicrobial compound (I) which is a peptide
CC having 8-50 amino acids, a net charge of 4, a hydrophobic moment (micron)
CC as a beta sheet which is 0.2 higher than its micro H as alpha helix, and
CC having detectable membrane disrupting activity against a microbial
CC pathogen, and substantially no membrane disrupting activity against
CC mammalian cells. (I) is useful for inhibiting microbial activity. (I)
CC has a detectable membrane disrupting activity against a pathogen, and is
CC useful for inhibiting non-microbial pathogenic activity also. (I) is also
CC useful for killing human sperm. The peptides are also provided in the
CC form of an expression vector comprising a nucleic acid encoding the
CC peptides. The peptides are useful for inhibiting the activity of

CC bacteria, and other microbial pathogens such as algae, fungi or protozoa
CC and for inhibiting non-microbial pathogens such as worms or arthropods,
CC and as spermicides for humans as the sperm membrane is atypical of human
CC cell membranes. (I) also has diagnostic uses e.g., in localizing an
CC infection or detecting sepsis. The peptides may act as binding molecules
CC and are useful to purify a target from blood, for qualitative or
CC quantitative analysis of analytes in in vitro sample, and for in vivo
CC imaging. Also, they are useful as molecular weight markers, as nutrient
CC source, as growth medium component for culturing microorganisms, as well
CC as a food ingredient for human consumption. The peptides have a greater
CC selectivity for bacterial versus mammalian lipids as compared to the
CC alpha helical peptides. Sequences AAG65536-47 represent amino acid
CC sequences of antimicrobial peptides.
XX
SQ Sequence 18 AA;
Query Match 41.4%; Score 67; DB 22; Length 18;
Best Local Similarity 47.1%; Pred. No. 0.028;
Matches 8; Conservative 9; Mismatches 0; Indels 0; Gaps 0;
QY 3 RRVRRRVVRRVRRV 19
Db 2 RIRRIIRIRIRIRI 18
RESULT 3
AAM06684
ID AAM06684 standard; peptide; 39 AA.
XX
AC AAM06684;
XX
DT 05-AUG-1997 (first entry)
XX
DE Protamine-like peptide analogue [+18RGD].
XX
KM Coagulation; anticoagulant; heparin; platelet aggregation; cell adhesion;
KM positively charged cluster; arginine; polycationic; decrease;
KM n-protamine; salmine protamine; protamine sulphate; salmon sperm.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1
FT Modified-site /note="acetylated"
FT Region 3..30
FT Binding-site /label=repeat_region
FT /note="4 tandem repeats of (Arg)2(Ala)2(Arg)2Ala motif"
FT /label=cell_adhesion_motif
FT Modified-site 39
FT /note="amidated"
XX
PN WO9635444-A1.
XX
PD 14-NOV-1996.
XX
PF 08-MAY-1996; 96WO-US06567.
XX
PR 08-MAY-1995; 95US-0436703.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Andrews PC, Stanley JC, Wakefield TW;
XX
DR WPI; 1997-011697/01.
XX
PT Peptide reversing the anticoagulant effects of heparin - is based on
PT protamine but has fewer positive charges for reduced toxicity
XX
PS Claim 31; Page 31; 42pp; English.
XX
CC Protamine sulphate (also called n-protamine or salmine protamine) is
CC a polycationic peptide derived from salmon sperm and is used to

OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
 OX NCBI_TaxID=134554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99350412; PubMed=10419946;
 RA Mündich L., Qiao X., Qiao J., Onodera S., Romantschuk M.,
 RA Hoegstraten D.;
 RT "Isolation of additional bacteriophages with genomes of segmented
 RT double-stranded RNA."
 RL J. Bacteriol. 181:4505-4508(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20473938; PubMed=11017801;
 RA Qiao X., Qiao J., Onodera S., Mündich L.;
 RT "Characterization of phi13, a bacteriophage related to phi6 and
 RT containing three dsRNA genomic segments."
 RL Virology 275:218-224(2000).
 DR EMBL; AF61667; AAC00439.1; -
 SQ SEQUENCE 230 AA; 24117 MW; 729BAAAF5566256 CRC64;

Query Match 43.2%; Score 70; DB 9; Length 230;
 Best Local Similarity 35.3%; Pred. No. 0.33;
 Matches 12; Conservative 15; Mismatches 7; Indels 0; Gaps 0;

QY 1 VRRVRRVVRVRRVRRVRRVRRVRRVRRV 34
 DB 11 VKRVVDVVVKVKKVVDLVVKAIKAIKIKIV 44

RESULT 3
 ID 071097 PRELIMINARY; PRT; 171 AA.
 AC 071097;
 DT 01-AUG-1998 (T-REMBLrel. 07, Created)
 DT 01-AUG-1998 (T-REMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (T-REMBLrel. 21, Last annotation update)
 DE PVI1.
 OS Bovine adenovirus type 3 (Mastadenovirus bos3).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10510;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MWR-1;
 RX MEDLINE=98105785; PubMed=9445040;
 RA Reddy P.S., Idamkant N., Zakharchouk A.N., Baxi M.K., Lee J.B.,
 RA Pyne C., Babluk L.A., Tikoo S.K.;
 RT "Nucleotide sequence, genome organization, and transcription map of
 RT bovine adenovirus type 3."
 RL J. Virol. 72:1394-1402(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MWR-1;
 RX MEDLINE=98318755; PubMed=9654686;
 RA Baxi M.K., Reddy P.S., Zakharchouk A.N., Idamkant N., Pyne C.,
 RA Babluk L.A., Tikoo S.K.;
 RT "Characterization of bovine adenovirus type 3 early region 2B."
 RL Virus Genes 16:313-316(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MWR-1;
 RX MEDLINE=98451815; PubMed=9778793;
 RA Lee J.B., Baxi M.K., Idamkant N., Reddy P.S., Zakharchouk A.N.,
 RA Pyne C., Babluk L.A., Tikoo S.K.;
 RT "Genetic organization and DNA sequence of early region 4 of bovine
 RT adenovirus type 3."
 RL Virus Genes 17:99-100(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MWR-1;
 RX MEDLINE=99119503; PubMed=9918888;
 RA Reddy P.S., Chen Y., Idamkant N., Pyne C., Babluk L.A., Tikoo S.K.;
 RT "Characterization of early region 1 and pIX of bovine adenovirus-3."
 RL Virology 253:299-308(1999).

RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MWR-1;
 RA Reddy P.S., Idamkant N., Zakharchouk A.N., Baxi M.K., Lee J.B.,
 RA Pyne C., Babluk L.A., Tikoo S.K.;
 RT Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF030154; AAD09724.1; -
 DR InterPro; IPR004912; Adeno_VII.
 DR Pfam; PF03228; Adeno_VII_1.
 SQ SEQUENCE 171 AA; 18959 MW; 0EACBIBC3C12519A CRC64;

Query Match 40.7%; Score 66; DB 12; Length 171;
 Best Local Similarity 54.3%; Pred. No. 0.73;
 Matches 19; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 2 RRVRRVVRVRRVRRVRRVRRVRRVRRVRR 36
 DB 84 RRRRRGVRVRRLRRSPRTALQRRVSVRRQV 118

RESULT 4
 ID 08TXS5 PRELIMINARY; PRT; 428 AA.
 AC 08TXS5;
 DT 01-JUN-2002 (T-REMBLrel. 21, Created)
 DT 01-JUN-2002 (T-REMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (T-REMBLrel. 21, Last annotation update)
 DE Uncharacterized protein.
 GN MK0585.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezhevaia K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malyn A.G., Koonin E.V., Kozayvkin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010352; AAM01800.1; -
 KM Complete genome.
 SQ SEQUENCE 428 AA; 48744 MW; 964D477CA264D13C CRC64;

Query Match 38.3%; Score 62; DB 17; Length 428;
 Best Local Similarity 45.7%; Pred. No. 5.1;
 Matches 16; Conservative 11; Mismatches 4; Indels 4; Gaps 2;

QY 3 RVRVRRVVRVRRVRRVRRVRRVRRVRRV 35
 DB 353 KYLQKVLREV-VRRVGTGRILVRIKRIARLYK 385

RESULT 5
 ID 0987V7 PRELIMINARY; PRT; 349 AA.
 AC 0987V7;
 DT 01-OCT-2001 (T-REMBLrel. 18, Created)
 DT 01-OCT-2001 (T-REMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (T-REMBLrel. 20, Last annotation update)
 DE Hypothetical protein ml16891.
 GN ML16891.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;

QY 7 RVVRVVRVVRVVRVVRVVRVVRV 34
 DB 828 RVLEVQAVYRAIRIPRIQGLRIIL 855

RESULT 15

RVVB_BRUME STANDARD; PRT; 346 AA.
 AC Q8Y15;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Holliday junction DNA helicase ruvb.
 GN RVVB OR EMEI0334.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kapratul V., Redkar R.J., Patra G., Mujer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-U.,
 RA Haselkorn R., Kyriades N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 CC -!- FUNCTION: The ruva-ruvb complex in the presence of ATP renatures
 cruciform structure in supercoiled DNA with palindromic sequence,
 indicating that it may promote strand exchange reactions in
 homologous recombination. RuvaB is an helicase that mediates the
 Holliday junction migration by localized denaturation and
 reannealing (By similarity).
 CC -!- SUBUNIT: Forms a complex with ruva (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE RVVB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: A809475; AAL51515.1; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_cent.
 DR InterPro; IPR004605; RVVB.
 DR Pfam; PF00004; AAA; 1.
 DR TIGRFAMs; TIGR00635; ruvb; 1.
 KM DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
 KW Complete proteome.
 FT NP_BIND 59 66 ATP (POTENTIAL).
 SQ SEQUENCE 346 AA; 38268 MW; F97710688F919FBC CRC64;

Query Match 31.2%; Score 50.5; DB 1; Length 346;
 Best Local Similarity 38.5%; Pred. No. 12;
 Matches 15; Conservative 5; Mismatches 10; Indels 9; Gaps 1;

QY 4 VVRVVRVVRVVRVVRVVRVVRV 33
 DB 189 IVRGRARIMQGISDGAREVARSCTPRIVGRLRRV 227

Search completed: June 9, 2003, 11:56:38
 Job time: 8.2766 secs

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
 RA Gilm S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Haelegh J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasaahara Y., Kleier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Maunda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nodack M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawara A., Oudaga B., Park S.H.,
 RA Perro V., Pohl T.M., Portelle D., Portoullis S., Prescott A.M.,
 RA Peseant E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Takakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 RA Toso V., Uchiyama S., Vandenbol M., Vannier P., Vassart A.,
 RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipet A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zuretein E., Yoshikawa H., Zanchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z99117; CAB1675.1; -.
 DR Subtilisin; B012696; udk.
 DR InterPro: IPR001324; PRK.
 DR InterPro: IPR00764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00478; PHRIKINASE.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRfams: TIGR00235; udk; 1.
 KW Transferase; Kinase; ATP-binding; Complete proteome.
 KW NP BIND 12 19 ATP (POTENTIAL).
 KW SEQUENCE 211 AA; 24487 MW; 6BFCF031D88C984C4 CRC64;
 SQ
 Query Match 31.5%; Score 51; DB 1; Length 211;
 Best Local Similarity 34.5%; Pred. No. 6.5;
 Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
 QY 3 RVRVRVVVRVVVRVVVRVVVRVVVR 31
 Db 138 RIRIRIMDRINRGSRISDVIEQYVVR 166
 RESULT 14
 ENV_HV1H3 STANDARD; PRT; 856 AA.
 AC P04624;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-1999 (Rel. 38, Last annotation update)
 DE Entolope polypeptide GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OC NCBI_taxid=11707;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85228248; PubMed=2988795;
 RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
 RA Shaw G.M., Wong-Saai F., Reddy E.P.;
 RT "HIV-1 env gene products synthesized in *E. coli* are recognized by
 RT antibodies present in the sera of AIDS patients.";
 RL Cell 41:979-986(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14100; AAA44679.1; -.
 DR HIV; M14100; ENVSHXB3.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 166 166
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CARBOHYD 356 356
 FT CARBOHYD 386 386
 FT CARBOHYD 392 392
 FT CARBOHYD 397 397
 FT CARBOHYD 406 406
 FT CARBOHYD 448 448
 FT CARBOHYD 463 463
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 624 624
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 750 750
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97188 MW; 3373C68B88C1AFC CRC64;
 Query Match 31.5%; Score 51; DB 1; Length 856;
 Best Local Similarity 32.1%; Pred. No. 26;
 Matches 9; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
 QY 3 RVRVRVVVRVVVRVVVRVVVRVVVR 31
 Db 138 RIRIRIMDRINRGSRISDVIEQYVVR 166

QY 12 VRRV--RRVRRVRRVRRVRRVRR 36
 ID YACG RHISN STANDARD; PRT; 305 AA.
 DB 14 LRRVYGRRTTRAI-RAVRRREFVRR 39

RESULT 8

QY 12 VRRV--RRVRRVRRVRRVRRVRR 36
 ID YACG RHISN STANDARD; PRT; 305 AA.
 DB 14 LRRVYGRRTTRAI-RAVRRREFVRR 39

RESULT 9

QY 12 VRRV--RRVRRVRRVRRVRRVRR 36
 ID YACG RHISN STANDARD; PRT; 305 AA.
 DB 14 LRRVYGRRTTRAI-RAVRRREFVRR 39

SEQUENCE FROM N.A.

QY 12 VRRV--RRVRRVRRVRRVRRVRR 36
 ID YACG RHISN STANDARD; PRT; 305 AA.
 DB 14 LRRVYGRRTTRAI-RAVRRREFVRR 39

RESULT 10

QY 12 VRRV--RRVRRVRRVRRVRRVRR 36
 ID YACG RHISN STANDARD; PRT; 305 AA.
 DB 14 LRRVYGRRTTRAI-RAVRRREFVRR 39

```

ID  PRT1 MUGCE  STANDARD;  PRT;  33 AA.
AC  P08130;
DT  01-AUG-1988 (Rel. 08, Created)
DT  01-AUG-1988 (Rel. 08, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Protamine M6/M7 (Mugil beta)
OS  Mugil cephalus (Flathead mullet) (Mugil japonicus).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC  Acanthomorpha; Acanthopterygii; Percomorpha; Mugilomorpha; Mugilidae;
OC  Mugil.
NCBI_TaxID=48193;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Sperm;
RX  MEDLINE=87279969; PubMed=3301825;
RA  Okamoto Y., Muta E., Ota S.;
RT  J. Biochem. 101:1017-1024(1987).
RL  J. Biochem. 101:1017-1024(1987).
CC  -1- FUNCTION: PROTIMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC  SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC  SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC  -1- SUBCELLULAR LOCATION: Nuclear.
CC  -1- TISSUE SPECIFICITY: TESTIS.
CC  -1- MISCELLANEOUS: THE SEQUENCE OF COMPONENT M6 IS SHOWN.
DR  PIR; A26762; A26762.
DR  PIR; B26762; B26762.
KW  Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW  Testis; DNA condensation; Nuclear protein.
FT  VARIANT 6 5 I -> Q (IN M7 COMPONENT).
FT  VARIANT 22 2 I -> M.
SQ  SEQUENCE 33 AA; 4473 MW; 4B407DE638A0D29E CRC64;

Query Match 33.3%; Score 54; DB 1; Length 33;
Best Local Similarity 51.9%; Pred. No. 0.45;
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 10 RYRVRRVRRVRRVRRVRRVRRVRR 36
DB 5 RETSRPIRRRRRRARAPIRRRRRVRR 31

RESULT 6
PRT2_CLUPA  STANDARD;  PRT;  31 AA.
AC  P02336;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  15-DEC-1998 (Rel. 37, Last annotation update)
DE  Protamine 2 (Clupeine 2).
OS  Clupea pallasii (Pacific herring), and
OS  Clupea harengus (Atlantic herring).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
OC  Clupea.
NCBI_TaxID=30724, 7950;
RN  [1]
RP  SEQUENCE.
RC  SPECIES=C.pallasii;
RX  MEDLINE=71157437; PubMed=5551645;
RA  Iwai K., Nakahara C., Ando T.;
RT  "Studies on protamines. XV. The complete amino acid sequence of the Z
RT  component of clupeine. Application of N leads to O acyl rearrangement
RT  and selective hydrolysis in sequence determination.";
RL  J. Biochem. 69:493-509(1971).
RN  [2]
RP  SEQUENCE.
RC  SPECIES=C.harengus;
RA  Chang W.J., Nukushina M., Ishii S., Nakahara C., Ando T.;
RT  Submitted (AUG-1970) to the PIR data bank.
RN  [3]
RX  X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
MEDLINE=92126280; PubMed=1772633;

```

```

RA  Baleschmidt P., Hansen F.B., Dodson E., Dodson G., Korber F.;
RT  "Structure of porcine insulin cocrystallized with clupeine Z.";
RL  Acta Crystallogr. B 47:975-986(1991).
CC  -1- FUNCTION: PROTIMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC  SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC  SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC  -1- SUBCELLULAR LOCATION: Nuclear.
CC  -1- TISSUE SPECIFICITY: TESTIS.
CC  -1- MISCELLANEOUS: CLUPEINE Z IS PROBABLY THE RESULT OF A CROSSOVER
CC  BETWEEN THE GENES FOR CLUPEINES YI AND YII.
DR  PIR; A02678; CLHRZ.
DR  PIR; A37576; CLHRZ.
DR  PDB; 7INS; 3I-TAN-94.
KW  Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW  Testis; DNA condensation; Nuclear protein; 3D-structure.
SQ  SEQUENCE 31 AA; 4165 MW; 092CB7F3AFC050 CRC64;

Query Match 33.0%; Score 53.5; DB 1; Length 31;
Best Local Similarity 54.3%; Pred. No. 0.48;
Matches 19; Conservative 0; Mismatches 9; Indels 7; Gaps 2;

QY 2 RYRVRRVRRVRRVRRVRRVRRVRRVRR 36
DB 3 RRSRRASRPVRR--RRPRVRR-----RRARR 30

RESULT 7
RL31_AERPE  STANDARD;  PRT;  105 AA.
AC  Q9YD25;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L31e.
GN  RPL31E OR APEL087.
OS  Aeropyrum pernix.
OC  Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC  Desulfurococcaceae; Aeropyrum.
OX  NCBI_TaxID=56636;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=KL1;
RX  MEDLINE=9310339; PubMed=10382966;
RA  Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA  Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anzai A., Kosugi H.,
RA  Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA  Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA  Yamazaki U., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA  Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT  "Complete genome sequence of an aerobic hyper-thermophilic
RT  crenarchaeon, Aeropyrum pernix KL1.";
RL  DNA Res. 6:83-101(1999).
CC  -1- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; AP000060; BAA80072.1; -
DR  InterPro; IPR000054; Ribosomal_L31e.
DR  Pfam; PF01198; Ribosomal_L31e; 1.
DR  PROSITE; PS01144; RIBOSOMAL_L31E; FALSE_NEG.
KW  Ribosomal protein; Complete proteome.
SQ  SEQUENCE 105 AA; 12527 MW; 7E5DF7999E74A098 CRC64;

Query Match 32.4%; Score 52.5; DB 1; Length 105;
Best Local Similarity 55.6%; Pred. No. 2.1;
Matches 15; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

```

```

RN NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C-125 / JCM 9153;
RX MEDLINE=20515582; PubMed=11056132;
RX Takami H., Nakaone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RX Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RX "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AP001511; BAB04994.1; -.
DR InterPro; IPR001324; PRK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; URIDINKINASE.
DR PRINTS; PRO0988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; udk; 1.
DR Transferase; Kinase; ATP-binding; Complete proteome.
DR NP BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 211 AA; 24387 MW; C2AFA2CB0030520B CRC64;

Query Match 34.6%; Score 56; DB 1; Length 211;
Best Local Similarity 37.9%; Pred. No. 1.6;
Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Oy 3 RVRRVRRVRRVRRVRRVRRVRRVRRVRR 31
Db 138 RIRRVADIRRGRTLESVLEQYTKVRR 166
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 3
ID1_CLUPEA STANDARD; PRT; 30 AA.
AC P02335;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Protamine YII (Clupeine YII).
OS Clupea pallasii (Pacific herring), and
OS Clupea harengus (Atlantic herring);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
OC Clupea.
OX NCBI_TaxID=30724, 7950;
RN [1]
RN SEQUENCE.
RP SPECIES=C.pallasii;
RX MEDLINE=73223106; PubMed=4664740;
RA Suzuki K., Ando T.;
RT "Studies on protamines. XVI. The complete amino acid sequence of
RT clupeine YII.";
RL J. Biochem. 72:1419-1432(1972).
RN [2]
RN SEQUENCE.
RP SPECIES=C.harengus;
RC Chang W.J., Nukushina M., Ishii S., Nakahara C., Ando T.;
RA Submitted (AUG-1970) to the PIR data bank.
CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF

```

```

CC CC Sperm during the haploid phase of spermatogenesis. They compact
CC sperm DNA into a highly condensed, stable and inactive complex.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: testis.
DR PIR; A02677; C1HRV2.
DR PIR; A37575; C1HR2A.
KW Chromosomal protein, Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
SQ SEQUENCE 30 AA; 4049 MW; 7F9BBB0FADAD566 CRC64;

Oy Query Match 34.3%; Score 55.5; DB 1; Length 30;
   Best Local Similarity 54.3%; Pred. No. 0.27;
   Matches 19; Conservative 0; Mismatches 9; Indels 7; Gaps 2;

Db 2 RRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
   ||| ||| ||| ||| ||| ||| |||
   2 RRRTRRASRPYRR--RRRRVRSR-----RRARRR 29

RESULT 4
ID_VCOT_ADE40 STANDARD; PRT; 185 AA.
VCOT_ADE40
AC Q89532;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
GN pVII.
OS Human adenovirus type 40.
OC Viruses; dsDNA viruses; no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxId=28284;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RX MEDLINE=94087748; PubMed=8263936;
RA Davison A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
RT "The DNA sequence of adenovirus type 40.";
RL J. Mol. Biol. 234:1308-1316(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RA Pieniazek N.J., Slemenda S.B., Pieniazek D., Luftig R.B.;
RC Submitted (FE8-1992) to the EMBL/GenBank/DBJ databases.
RL -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).
CC -----
DR EMBL; L19443; AAC13963.1; -.
DR EMBL; M86655; AAA42526.1; -.
DR InterPro; IPR004912; Adeno_VII.
DR Pfam; PF03228; Adeno_VII; 1.
DR PROPEP 1 BY SIMILARITY.
FT CHAIN 24 185 MAJOR CORE PROTEIN.
FT SITE 23 24 CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT SITE (POTENTIAL)
SQ SEQUENCE 185 AA; 20518 MW; 4FBBOB535FF218A9B CRC64;

Oy Query Match 34.0%; Score 55; DB 1; Length 185;
   Best Local Similarity 39.0%; Pred. No. 1.9;
   Matches 16; Conservative 5; Mismatches 10; Indels 10; Gaps 1;

Db 6 RRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36
   ||| ||| ||| ||| ||| ||| |||
   92 RRVARRKRRRLQRRRRRPTAAMTAAARAVLRRAORIGRRAMR 132

RESULT 5
PRTB_MUGOE

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:40:31 ; Search time 7.2766 Seconds

(without alignments)
205.199 Million cell updates/sec

Title: US-10-079-075-6

Perfect score: 162

Sequence: 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR

Scoring table: BLOSUM62

Searched: 11892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 11892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: SwisProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	34.6	193	VC07_ADE04	O96831 human adeno
2	56	34.6	211	URK_BACD	O96831 human adeno
3	55.5	34.3	30	PRT1_CUPE	P02335 clupea pall
4	55	34.0	185	VC07_ADE40	O96831 human adeno
5	54	33.3	33	PRTB_MOCCE	P08130 mugil cepha
6	53.5	33.0	31	PRT2_CUPE	P02336 clupea pall
7	52.5	32.4	105	RL31_AERSE	O95485 aeropyrum p
8	52.5	32.4	305	Y4CG_RHNS	P85389 rhizobium s
9	52	32.1	304	NADA_MERTH	O27855 methanobact
10	51	31.5	32	PRT_ORYLA	O91185 oryza lat
11	51	31.5	100	HSP2_ALOSE	P35312 aloatata se
12	51	31.5	183	VC07_ADEB2	O96624 bovine aden
13	51	31.5	211	URK_BACSU	O32033 bacillus su
14	51	31.5	86	ENV_HV1H3	P04624 human immun
15	50.5	31.2	346	RUVB_BRUME	O89145 bruceella me
16	50	30.9	99	VH5B_BPT7	P03751 bacterioph
17	50	30.9	263	IF2A_SULTO	O97380 sulfolobus
18	50	30.9	451	PSS_ECOLI	P23830 escherichia
19	49	30.2	170	YACP_BACSU	P37574 bacillus su
20	49	30.2	266	IF2A_SULTO	O97279 sulfolobus
21	49	30.2	378	GB12_MOUSE	P27600 mus musculu
22	49	30.2	378	GB12_MOUSE	O63210 rat mus musculu
23	49	30.2	380	GB12_MOUSE	O03113 homo sapien
24	49	30.2	399	YI99_ARCFU	O28380 archaeoglob
25	49	30.2	853	ENV_HV1M1	P19551 human immun
26	49	30.2	856	ENV_HV1M1	P03375 human immun
27	49	30.2	856	ENV_HV1M2	P04578 human immun
28	49	30.2	856	ENV_HV1M2	P070626 human immun
29	49	30.2	856	ENV_HV1M2	P03376 human immun
30	49	30.2	856	ENV_HV1M2	P03377 human immun
31	48.5	29.9	34	PRT2_THUTH	P02332 thunnus thy
32	48	29.6	58	HSP2_HORSE	P15343 equus cabal
33	48	29.6	62	HSP2_HORSE	P15342 equus cabal

34	48	29.6	70	1	L2M1_ADEB2	O96626 bovine aden
35	48	29.6	85	1	Y02A_BPT4	P39231 bacterioph
36	48	29.6	846	1	ENV_HV1M1	P18799 human immun
37	48	29.6	847	1	ENV_HV1M2	P05880 human immun
38	48	29.6	856	1	ENV_HV1M1	P31872 human immun
39	47.5	29.3	102	1	HSP2_PANTR	P35100 pan troglod
40	47.5	29.3	317	1	RPSD_STRAU	P27785 streptomyc
41	47.5	29.3	336	1	HRD1_STRCO	P18182 streptomyc
42	47	29.0	47	1	HSP1_CAVPO	P35304 cavia porce
43	47	29.0	208	1	URK_CLOPE	O8xj16 clostridium
44	47	29.0	851	1	ENV_HV1B8	P04582 human immun
45	47	29.0	863	1	ENV_HV1Z8	P05882 human immun

ALIGNMENTS

RESULT 1
VC07_ADE04 STANDARD; PRT; 193 AA.
ID VC07_ADE04
AC O96831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
GN EVII.
OS Human adenovirus type 4.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI-6;
RA Tatarskii L., Szawlowski P.W.S., McRay J., Russell W.C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb.ch).
CC
CC EMBL; U70921; AAC83411.1; -
CC InterPro; IPR004912; Adeno VII.
CC Pfam; PF03228; Adeno VII; I.
DR PROPEP
FT CHAIN 1 24 BY SIMILARITY.
FT SITE 25 193 CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT SITE 24 25 (POTENTIAL).
SQ SEQUENCE 193 AA; 21358 MW; 43137E07DB379DD0 CRC64;
Query Match 34.6%; Score 56; DB 1; Length 193;
Best Local Similarity 40.5%; Pred. No. 1.5;
Matches 15; Conservative 6; Mismatches 10; Indels 6; Gaps 2;
Oy 2 RRVRR-----VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR
Db 103 RRIARRHRTTAMRAARALIRRAARRGRAMRAARR 139
RESULT 2
URK_BACD STANDARD; PRT; 211 AA.
ID URK_BACD
AC O9KD8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
GN UDK OR BH1275.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

	Matches	11; Conservative	8; Mismatches	10; Indels	0; Gaps	0; Ns
QY	3	RVRRRVAVRVRVRVRVRVRVRVR	31			
Db	138	RIIRMTDIREGRITLESVIEQYIKVR	166			

RESULT 3
CLHRV2
protamine YII - Pacific herring

C:Species: *Clupea pallasii*, *Clupea harengus pallasii* (Pacific herring)
C:Date: 12-Aug-1981 #sequence_revision 12-Aug-1981 #text_change 16-Feb-1997
C:Accession: A38052; A02877
R: Suzuki, K.; Ando, T.
J. Biochem. 72, 1419-1432, 1972
A: Title: Studies on proclamines. XVI. The complete amino acid sequence of clupeine VII
A: Reference number: A38052; MUID: 73223106; PMID: 4664740
A: Accession: A38052
A: Molecule type: protein
A: Residues: 1-30 <SUZ>
C: Superfamily: prolamine Y2
C: Keywords: chromosomal protein, DNA binding, spermatogenesis

	34.3%	Score 55.5;	DB 1;	Length 30;
Query Match %	54.3%	Pred No. 1.7;		
Best Local Similarity	19;	Conservative	0;	Mismatches 9; Indels 7; Gaps 2
Matches	2 RRVVRRVVVRVVVRVVVRVVVRVVRR	36		
Oy	2 RRVVRRVVVRVVVRVVVRVVVRVVRR	36		
Dd	2 RRRTTRRASPVR--RRPRPSRR-----RRARR	29		

```

RESULT 4
CLHR2A
protamine y7i - Atlantic herring
M.Alternate names: clupeine
C.Species: Clupea harengus, Clupea harengus harengus (Atlantic herring)
C.Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Feb-1997
C.Accession: A37575, A02677
R.Chang, W.J.; Nukushina, M.; Ishii, S.; Nakahara, C.; Ando, T.
submitted to the Atlas, August 1970
A.Reference number: A37575
A.Accession: A37575
A.Molecule type: protein
A.Residues: 1-30 <CHA>
C.Superfamily: protamine Y2
C.Keywords: chromosomal protein; DNA binding; spermatogenesis

Query Match          34.3%   Score 55.5; DB 1; Length 30;
Best Local Similarity 54.3%   Pred. No. 1.7;
Matches 19; Conservative 0; Mismatches 9; Indels 7; Gaps 2

OY      2  RRVRVRVRVRVRVRVRVRVRVRVRVRVRVRVR 36
DB      2  RRRTRRASPVR--RRPRRSRR-----RRRR 29

```

```

RESULT 5
T43950
hypothetical protein LTL [imported] - human herpesvirus 6
C:Species: human herpesvirus 6
A:Variety: strain HST
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T43950 #sequence_revision 214057
R:Issigawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kaw
U:Virol. 73, 8053-8063, 1999
A>Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and
A:Reference number: 222732; NUID:99412319; PMID:10482554
A:Accession: T43950
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-115 <ISE>

```

	A1.Cros8-references: EMBL:AB021506; NID:g94995977; PID:BAA7821.1.1; PID:g94995978
	A1.Experimental source: strain HSTr, pop. variant B
	A1.Accession: T44057
	A1.Status: preliminary; translated from GB/EMBL/DDBJ
	A1.Molecule type: DNA
	A1.Residues: 1-115 <IS2>
	A1.Cross-references: EMBL:AB021506; NID:g94995977; PID:BAA7821.1.1; PID:g94995978
	A1.Experimental source: strain HSTr, pop. variant B
	C1.Genetics:
	A1.Gene: LTL; RJ1
Query Match	34.3%; Score 55.5; DB 2; length 115;
Best Local Similarity	57.9%; Pred. No. 5.9;
Matches 22; Conservative	4; Mismatches 9; Indels 3; Gaps 3
Cy	1 VRVRRVRVVR-VRRVR-RVRVRVVRVVR-VRRVR 35 : : : : : : : : 14 VRAVRVRVRVRVRVRVRVRVRVRVRVRVRVRVR 51
Db	

RESULT 6
S24610
major membrane cytoskeletal protein artticultin, 86k - *Euglena gracilis*
C:Species: *Euglena gracilis*
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Sep-1997
C:Accession: B43417; S24610
R:Maris, J.A.; Bouck, G.B.
J. Cell Biol. 118, 1465-1475, 1992
A:Title: The two major membrane skeletal proteins (artticultins) of *Euglena gracilis* defi
A:Reference number: A43417; MUID:92394973; PMID:1522117
A:Accession: B43417
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-650 <MAR>
A:Cross-references: EMBL:Z13963; NID:G18405; PID:G18406

	Query Match	33.6%;	Score 54.5;	DB 2,	length 650;
	Best Local Similarity	48.5%;	Pred. No. 37;		
	Matches	16;	Conservative	7;	Mismatches
				9;	Indels
				1;	Gaps
QY	2 RRVVRRVVRRVVRR-RVVRRVVRRVVRRV	33			
	: : :				
Db	201 RRVPVEQIVRRVQVVERLVKCVVARRHQV	233			

RESULT 7
 S24609
 cytoskeletal protein - *Euglena gracilis*
 C:Species: *Euglena gracilis*
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
 C:Accession: S24609; A43417
 R:Bouck, G.B.
 submitted to the EMBL Data Library, June 1992
 A:Reference number: S24609
 A:Accession: S24609
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-651 <BOU>
 A:Cross-references: EMBL:Z13962; NID:g18403; PID:g18404
 R:Marx, J.A.; Bouck, G.B.
 J. Cell Biol. 118, 1465-1475, 1992
 A:Title: The two major membrane skeletal proteins (articulins) of *Euglena gracilis* defined
 A:Reference number: A43417; MUID:92394973; PMID:1522117
 A:Accession: A43417
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-145, 'T', 147-187, 'T', 189-348, 'T', 350-651 <MAR>
 A>Note: sequence extracted from NCBI backbone (NCBI:113032)

Query Match	33.6%	Score	54.5	DB	2	Length	651
Best Local Similarity	40.9%	Pred. No.	37				
Matches	18	Conservative	8	Mismatches	9	Indels	9
						Gaps	3

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:48:06 ; Search time 14.9362 Seconds
(without alignments)
231.709 Million cell updates/sec

Title: US-10-079-075-6

Perfect score: 162
Sequence: 1 VRRVRRVRRVRRVRRVRRVRRVRRVRRVRR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.5	36.7	219	C84477	hypothetical prote
2	56	34.6	241	C83809	uridine kinase udk
3	55.5	34.3	30	1 CLHRY2	protamine YII - Pa
4	55.5	34.3	30	1 CLHRY2	protamine YII - Ac
5	55.5	34.3	115	2 T43950	hypothetical prote
6	54.5	33.6	650	S24610	major membrane cyt
7	54.5	33.6	651	S24609	cytoskeletal prote
8	54	33.3	33	2 A26762	protamine (mugilin
9	54	33.3	33	2 B26762	protamine (mugilin
10	54	33.3	1153	2 F84468	hypothetical prote
11	53.5	33.0	31	1 CLHRY2	protamine Z - Pacl
12	53.5	33.0	31	1 CLHRY2	protamine Z - Atla
13	53	32.7	114	2 T17699	arginine-rich prot
14	53	32.7	114	2 F84500	hypothetical prote
15	53	32.7	1218	2 E84537	hypothetical prote
16	52.5	32.4	105	2 H72708	hypothetical prote
17	52	32.1	304	2 C69111	probable ribosomal
18	51.5	31.8	34	2 JN0582	quintolinate synthet
19	51.5	31.8	170	2 JN0588	protamine (secombr
20	51	31.5	32	2 F51089	hypothetical 20K p
21	51	31.5	100	2 S33338	protamine - Japane
22	51	31.5	180	2 F85355	protamine p2 - red
23	51	31.5	211	2 G69728	hypothetical prote
24	51	31.5	233	2 G95987	uridine kinase udk
25	51	31.5	233	2 C72784	probable two-compo
26	50.5	31.2	346	2 AH3293	hypothetical prote
27	50	30.9	99	1 QSBP87	holliday junction
28	50	30.9	240	2 F82790	host specificity p
29	50	30.9	451	2 AE0831	GMP synthase XF056
30	50	30.9	451	2 AE0831	CDPdiacylglycerol-

30	50	30.9	452	1	H65036	CDPdiacylglycerol-
31	50	30.9	452	2	AG0397	CDPdiacylglycerol-
32	50	30.9	452	2	B85905	CDPdiacylglycerol-
33	50	30.9	452	2	D91060	CDPdiacylglycerol-
34	49.5	30.6	763	2	T24367	hypothetical prote
35	49.5	30.6	814	2	T30950	hypothetical prote
36	49.5	30.6	829	2	E87757	hypothetical prote
37	49.5	30.6	2712	2	T30949	protein C44B4.1b
38	49.5	30.6	3864	2	D87757	hypothetical prote
39	49	30.2	166	2	D87664	protein C44B4.1a
40	49	30.2	184	2	S66125	conserved hypochet
41	49	30.2	170	2	T13097	probable minor tal
42	49	30.2	255	2	T28054	hypothetical prote
43	49	30.2	266	2	C90257	translation initia
44	49	30.2	297	2	H72530	hypothetical prote
45	49	30.2	379	2	A41095	GTP-binding regula

ALIGNMENTS

RESULT 1
C84477
hypothetical protein AC2906420 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #ext_change 02-Feb-2001
C/Accession: C84477
R/lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; NCID:20083487; PMID:10617197
A/Accession: C84477
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-349 <STO>
A/Cross-references: GB:AE002093; NID:G4646226; PIDN:AAD36890.1; GSPDB:GN00139
C/Genetics:
A/Map position: 2

Query Match 36.7%; Score 59.5; DB 2; Length 349;
Best Local Similarity 41.7%; Pred. No. 5.8;
Matches 15; Conservative 8; Mismatches 12; Indels 1; Gaps 1;

QY 2 RRVRR-VRRVRRVRRVRRVRRVRRVRRVRR 36
DB 274 RRTKRRVIRRRVIRRRVIRRRVIRRRVIRRR 309

RESULT 2
C83809
uridine kinase udk [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #ext_change 15-Jun-2001
C/Accession: C83809
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hlt
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; NCID:20512582; PMID:11058132
A/Accession: C83809
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-211 <STO>
A/Cross-references: GB:AP001511; GB:BA000004; NID:G10173727; PIDN:BA04994.1; GSPDB:GNO
A/Experimental source: strain C-125
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

Query Match 34.6%; Score 56; DB 2; Length 211;
Best Local Similarity 37.9%; Pred. No. 9;

TELEPHONE: 212-705-5000
 TELEFAX: 212-765-2519
 INFORMATION FOR SEQ ID NO: 54:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 21 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 5945507e
 US-08-932-682-54

Query Match 45.0%; Score 49; DB 2; Length 21;
 Best Local Similarity 44.4%; Pred. NO. 0.41;
 Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 6 RVRVRVRVRVRVRVRVR 23
 |:|:|:|:|:|:|:|:|:|
 Db 1 RIRRPRIIRICIGQVVR 18

Search completed: June 9, 2003, 12:05:06
 Job time : 23.1702 secs

Db 1 RRIYRAIRHPRIRRCIGVRR 24

RESULT 13

US-08-786-748A-54
Sequence 54, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-54
Query Match 45.0%; Score 49; DB 1; Length 21;
Best Local Similarity 44.4%; Pred. No. 0.41;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 6 RRVRRVRRVRRVRRV 23
Db 1 RIRPRIRIRICIGVRR 18
RESULT 14
US-08-786-748A-55
Sequence 55, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-55
Query Match 45.0%; Score 49; DB 1; Length 21;
Best Local Similarity 47.1%; Pred. No. 0.41;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 6 RRVRRVRRVRRVRRV 22
Db 1 RIRPRIRIRICIGVRR 17
RESULT 15
US-08-932-682-54
Sequence 54, Application US/08932682
Patent No. 5945507
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BORTS, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,682
FILING DATE: 18-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/786,748
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421-A
TELECOMMUNICATION INFORMATION:

DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-09-340-154-64

Query Match 47.7%; Score 52; DB 3; Length 27;
Best Local Similarity 25.0%; Pred. No. 0.21;
Matches 6; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRRVRRVRRVRR 24
DB 3 RKIKRIKKIVRKFRIRIALIKRK 26

RESULT 7
US-09-482-611B-64
Sequence 64, Application US/09482611B
Patent No. 6448391
GENERAL INFORMATION:
APPLICANT: Garbano, Joan
TITLE OF INVENTION: Ubiquitin-Lytic Peptide Fusion Gene Constructs, Protein Products
TITLE OF INVENTION: Therefrom, and Methods of Making and Using Same
FILE REFERENCE: 2093-149
CURRENT APPLICATION NUMBER: US/09/482,611B
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: US 08/801,028
PRIOR FILING DATE: 1997-02-19
PRIOR APPLICATION NUMBER: US 08/279,472
PRIOR FILING DATE: 1994-07-22
NUMBER OF SEQ ID NOS: 102
SOFTWARE: Patentin version 3.1
SEQ ID NO 64
LENGTH: 27
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Lytic Peptide
US-09-482-611B-64

Query Match 47.7%; Score 52; DB 4; Length 27;
Best Local Similarity 25.0%; Pred. No. 0.21;
Matches 6; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRRVRRVRRVRR 24
DB 3 RKIKRIKKIVRKFRIRIALIKRK 26

RESULT 8
PCT-US95-09338-64
Sequence 64, Application PC/TUS9509338
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
NUMBER OF SEQUENCES: 98
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09338
FILING DATE: 21-JUL-1994
PRIOR APPLICATION DATA: 08/279,472
APPLICATION NUMBER: 08/279,472
FILING DATE: 22-JUL-1994
INFORMATION FOR SEQ ID NO: 64:

SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US95-09338-64

Query Match 47.7%; Score 52; DB 5; Length 27;
Best Local Similarity 25.0%; Pred. No. 0.21;
Matches 6; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRRVRRVRRVRR 24
DB 3 RKIKRIKKIVRKFRIRIALIKRK 26

RESULT 9
PCT-US95-09339-64
Sequence 64, Application PC/TUS9509339
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
NUMBER OF SEQUENCES: 98
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09339
FILING DATE: 21-JUL-1994
PRIOR APPLICATION DATA: 08/279,472
APPLICATION NUMBER: 08/279,472
FILING DATE: 22-JUL-1994
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US95-09339-64

Query Match 47.7%; Score 52; DB 5; Length 27;
Best Local Similarity 25.0%; Pred. No. 0.21;
Matches 6; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRRVRRVRRVRR 24
DB 3 RKIKRIKKIVRKFRIRIALIKRK 26

RESULT 10
US-08-179-632-9
Sequence 9, Application US/08179632
Patent No. 5607914
GENERAL INFORMATION:
APPLICANT: Rao, A. Gururaj; Zhong, Lingxiu
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES

? FILING DATE: 22-JUL-1994
 ? CLASSIFICATION: 536
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: WALKER, BARBARA W.
 ? REGISTRATION NUMBER: 35,400
 ? REFERENCE/DOCKET NUMBER: 2093-117A
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (202)783-6040
 ? TELEFAX: (202)783-6031
 ? INFORMATION FOR SEQ ID NO: 64:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 27
 ? TYPE: AMINO ACID
 ? TOPOLOGY: LINEAR
 ? MOLECULE TYPE: PEPTIDE
 ? DESCRIPTION: NO
 ? HYPOTHETICAL: NO
 ? FRAGMENT TYPE: COMPLETE PEPTIDE
 ? ORIGINAL SOURCE: SYNTHETIC
 ? IMMEDIATE SOURCE: SYNTHETIC
 ? PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
 ?

Query Match	47.7%	Score 52;	DB 2;	Length 27;
Best Local Similarity	25.0%	Pred. No. 0.21;		
Matches 6;	Conservative 13;	Mismatches 5;	Indels 0;	Gaps 0

QY 1 RRVRRVRVRRVRRVRRVRR 24
|:::|:::|:::|:
Db 3 RKILRKIKIVRKFIRIALIKRK 26

RESULT 5
 US-08-801-028-64
 : Sequence 64, Application US/08801028
 : Patent No. 6018102
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: JOAN GARBARINO
 : APPLICANT: JESSE M. JAYNES
 : APPLICANT: WILLIAM BEKNAP
 : TITLE OF INVENTION: USQUITTIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUCT
 :
 : NUMBER OF SEQUENCES: 98
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: STEVEN J. HULTQUIST
 : ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
 : STREET: 200 PARK DRIVE, SUITE 210
 : STREET: P.O. BOX 14329
 : CITY: RESEARCH TRIANGLE PARK
 : STATE: NORTH CAROLINA
 : COUNTRY: USA
 : ZIP: 27709
 :
 : COMPUTER READABLE FORM:
 :
 : MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
 : COMPUTER: APPLE MACINTOSH
 : OPERATING SYSTEM: MACINTOSH
 : SOFTWARE: M.S. WORD 5.0
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/801,028
 : FILING DATE: 19-FEB-1997
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/279,472
 : FILING DATE: JULY 22, 1994
 : APPLICATION NUMBER: 08/225,476
 : FILING DATE: 04-20-94
 : APPLICATION NUMBER: 08/225,476
 : FILING DATE: 04-08-94
 : APPLICATION NUMBER: 08/039,620
 : FILING DATE: 06-04-93
 : APPLICATION NUMBER: 08/148,491
 : FILING DATE: 11-08-93
 : APPLICATION NUMBER: 08/148,889
 : FILING DATE: 11-08-93
 :
 : ATTORNEY/AGENT INFORMATION:
 :

NAME: MASSEMAN, FRAN S.
REGISTRATION NUMBER: 34,273
REFERENCE/DOCKET NUMBER: 4013-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 990-9531
TELEFAX: (919) 990-9532
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: NO
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED

Query Match	47.7%	Score	52	DB	3	Length	27
Best Local Similarity	25.0%	Pred. NC	0.21				
Matches	6	Conservative	13	Mismatches	5	Indels	0
						Gaps	0

```

QY      1 RRVRVRVRVRVRVRVRVRVR 24
        |:::|:::|:::|:::|:::
DB      3 RKILRKIKIVRKFIKIALIKRK 26

```

RESULT 6
 US-09-340-154-64
 : Sequence 64, Application US/09340154
 : Patent No. 6084156
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Jesse M. Jaynes
 :
 : TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
 :
 : TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
 : TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
 :
 : NUMBER OF SEQUENCES: 98
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
 : STREET: 555 Thirteenth Street N.W.
 :
 : CITY: Washington
 :
 : STATE: D. C.
 :
 : COUNTRY: USA
 :
 : ZIP: 20004
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
 :
 : COMPUTER: IBM COMPATIBLE
 :
 : OPERATING SYSTEM: DOS
 :
 : SOFTWARE: Wordperfect 5.1+
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/340,154
 :
 : FILING DATE:
 :
 : CLASSIFICATION:
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/505,486
 :
 : FILING DATE: 21-JUL-1995
 :
 : APPLICATION NUMBER: U.S. 08/279,472
 :
 : FILING DATE: 22-JUL-1994
 :
 : ATTORNEY/AGENT INFORMATION:
 :
 : NAME: WALKER, BARBARA W.
 :
 : REGISTRATION NUMBER: 35,400
 :
 : REFERENCE/DOCKET NUMBER: 2093-117A
 :
 : TELECOMMUNICATION INFORMATION:
 :
 : TELEPHONE: (202) 783-6040
 :
 : TELEFAX: (202) 783-6031
 :
 : INFORMATION FOR SEQ ID NO: 64:
 :
 : SEQUENCE CHARACTERISTICS:
 :
 : LENGTH: 27
 :
 : TYPE: AMINO ACID
 :
 : TOPOLOGY: LINEAR
 :
 : MOLECULAR TYPE:

TITLE OF INVENTION: WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 6601 Woodward Avenue
CITY: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44mb, 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6;
SOFTWARE: ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7MK-060548-00233
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-17

Query Match 49.5%; Score 54; DB 2; Length 38;
Best Local Similarity 52.2%; Pred. No. 0.16; Mismatches 11; Indels 0; Gaps 0;

QY 1 RRVRRRVRRVRRVRRVRRVRR 23
DB 3 RRAARRRARRARRARRARRARR 25

RESULT 3
US-08-436-703B-5
Sequence 5, Application US/08436703B
Patent No. 5919761
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR
TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
TITLE OF INVENTION: WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 6601 Woodward Avenue
CITY: Detroit
STATE: Michigan

COUNTRY: United States of America
ZIP: 48226
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 1.44mb, 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6;
SOFTWARE: ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,703B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: 7MK-060548-00233
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-5

Query Match 49.5%; Score 54; DB 2; Length 39;
Best Local Similarity 52.2%; Pred. No. 0.16; Mismatches 12; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 RRVRRRVRRVRRVRRVRRVRR 23
DB 3 RRAARRRARRARRARRARRARR 25

RESULT 4
US-08-505-486-64
Sequence 64, Application US/08505486
Patent No. 5955573
GENERAL INFORMATION:
APPLICANT: Jesse M. Jaynes
TITLE OF INVENTION: URICOTIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,486
FILING DATE: 21-JUL-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/279,472

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:50:31 ; Search time 8.17021 Seconds
(without alignments)
86,430 Million cell updates/sec

Title: US-10-079-075-5
Perfect score: 109
Sequence: 1 RRVRRVRRVRRVRRVRRVRR 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.5	52.8	882	4 US-09-413-814-78	Sequence 78, Appl
2	54	49.5	38	2 US-08-436-703B-17	Sequence 17, Appl
3	54	49.5	39	2 US-08-436-703B-5	Sequence 5, Appl
4	52	47.7	27	2 US-08-505-486-64	Sequence 64, Appl
5	52	47.7	27	3 US-08-801-028-64	Sequence 64, Appl
6	52	47.7	27	3 US-09-340-154-64	Sequence 64, Appl
7	52	47.7	27	4 US-09-482-611B-64	Sequence 64, Appl
8	52	47.7	27	5 PCT-US95-09338-64	Sequence 64, Appl
9	52	47.7	27	5 PCT-US95-09339-64	Sequence 64, Appl
10	51	46.8	31	1 US-08-179-632-9	Sequence 9, Appl
11	51	46.8	31	1 US-08-440-174A-9	Sequence 9, Appl
12	51	46.8	31	5 PCT-US95-00062-9	Sequence 9, Appl
13	49	45.0	21	1 US-08-786-748A-54	Sequence 54, Appl
14	49	45.0	21	1 US-08-786-748A-55	Sequence 55, Appl
15	49	45.0	21	2 US-08-932-682-54	Sequence 54, Appl
16	49	45.0	21	2 US-08-932-682-55	Sequence 55, Appl
17	49	45.0	23	2 US-08-505-486-63	Sequence 63, Appl
18	49	45.0	23	3 US-08-801-028-63	Sequence 63, Appl
19	49	45.0	23	3 US-09-340-154-63	Sequence 63, Appl
20	49	45.0	23	4 US-09-482-611B-63	Sequence 63, Appl
21	49	45.0	23	5 PCT-US95-09338-63	Sequence 63, Appl
22	49	45.0	23	5 PCT-US95-09339-63	Sequence 63, Appl
23	49	45.0	24	1 US-08-786-748A-57	Sequence 57, Appl
24	49	45.0	24	2 US-08-932-682-57	Sequence 57, Appl
25	49	45.0	27	2 US-08-505-486-65	Sequence 65, Appl
26	49	45.0	27	2 US-08-801-028-65	Sequence 65, Appl
27	49	45.0	27	3 US-09-340-154-65	Sequence 65, Appl

28	49	45.0	27	4 US-09-482-611B-65	Sequence 65, Appl
29	49	45.0	27	5 PCT-US95-09338-65	Sequence 65, Appl
30	49	45.0	27	5 PCT-US95-09339-65	Sequence 65, Appl
31	47	43.1	24	1 US-08-786-748A-60	Sequence 60, Appl
32	47	43.1	24	1 US-08-786-748A-62	Sequence 62, Appl
33	47	43.1	24	2 US-08-932-682-60	Sequence 60, Appl
34	47	43.1	24	2 US-08-932-682-62	Sequence 62, Appl
35	47	43.1	28	1 US-08-786-748A-43	Sequence 43, Appl
36	47	43.1	28	1 US-08-786-748A-46	Sequence 46, Appl
37	47	43.1	28	1 US-08-786-748A-48	Sequence 48, Appl
38	47	43.1	28	2 US-08-932-682-43	Sequence 43, Appl
39	47	43.1	28	2 US-08-932-682-46	Sequence 46, Appl
40	47	43.1	28	2 US-08-932-682-48	Sequence 48, Appl
41	46	42.2	21	1 US-08-786-748A-50	Sequence 50, Appl
42	46	42.2	21	1 US-08-786-748A-53	Sequence 53, Appl
43	46	42.2	21	2 US-08-932-682-50	Sequence 50, Appl
44	46	42.2	21	2 US-08-932-682-53	Sequence 53, Appl
45	46	42.2	24	1 US-08-786-748A-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-09-413-814-78
; Sequence 78, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bioecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hottle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolysaccharide compounds
; FILE REFERENCE: PCT/US 99/2335
; CURRENT APPLICATION NUMBER: US/09/413,814
; EARLIER FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 78
; TYPE: PRT
; LENGTH: 882
; ORGANISM: Sorangium cellulosum
US-09-413-814-78
Query Match 52.8%; Score 57.5; DB 4; Length 882;
Best Local Similarity 60.0%; Pred. No. 1.2;
Matches 15; Conservative 3; Mismatches 6; Indels 1; Gaps 1;
QY 1 RRVRRVRRVRRVRRVRRVRR 24
DB 601 RRVRRVRRVRRVRRVRRVRR 625
US-08-436-703B-17
; Sequence 17, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR

```
; Sequence 7, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-7
```

```
Query Match          100.0%; Score 109; DB 1; Length 42;
Best Local Similarity 100.0%; Pred. No. 8e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 RRVRRRRRRRRRRRRRRRRRRR 24
DB 1 RRVRRRRRRRRRRRRRRRRRRR 24
```

```
RESULT 13
US-09-785-058-7
; Sequence 7, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-7
```

```
Query Match          100.0%; Score 109; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 8e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 RRVRRRRRRRRRRRRRRRRRRR 24
DB 1 RRVRRRRRRRRRRRRRRRRRRR 24
```

```
RESULT 14
US-09-785-059-7
; Sequence 7, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A35377 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
```

```
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-7
```

```
Query Match          100.0%; Score 109; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 8e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 RRVRRRRRRRRRRRRRRRRRRR 24
DB 1 RRVRRRRRRRRRRRRRRRRRRR 24
```

```
RESULT 15
US-10-079-075-7
; Sequence 7, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-7
```

```
Query Match          100.0%; Score 109; DB 24; Length 42;
Best Local Similarity 100.0%; Pred. No. 8e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 RRVRRRRRRRRRRRRRRRRRRR 24
DB 1 RRVRRRRRRRRRRRRRRRRRRR 24
```

```
Search completed: June 9, 2003, 12:25:18
Job time : 104.915 secs
```

```

PCT-US02-04812-6
; Sequence 6, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-6

```

```

Query Match          100.0%; Score 109; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 6, 7e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RRVRRRRRRRRRRRRRRRRRRRRR 24
Db 13 RRVRRRRRRRRRRRRRRRRRRRRR 36

```

```

RESULT 8
US-09-785-058-6
; Sequence 6, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-6

```

```

Query Match          100.0%; Score 109; DB 21; Length 36;
Best Local Similarity 100.0%; Pred. No. 6, 7e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RRVRRRRRRRRRRRRRRRRRRRRR 24
Db 13 RRVRRRRRRRRRRRRRRRRRRRRR 36

```

```

RESULT 9
US-09-785-059-6
; Sequence 6, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-6

```

```

Query Match          100.0%; Score 109; DB 21; Length 36;
Best Local Similarity 100.0%; Pred. No. 6, 7e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RRVRRRRRRRRRRRRRRRRRRRRR 24
Db 13 RRVRRRRRRRRRRRRRRRRRRRRR 36

```

```

RESULT 10
US-10-079-075-6
; Sequence 6, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-6

```

```

Query Match          100.0%; Score 109; DB 24; Length 36;
Best Local Similarity 100.0%; Pred. No. 6, 7e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RRVRRRRRRRRRRRRRRRRRRRRR 24
Db 13 RRVRRRRRRRRRRRRRRRRRRRRR 36

```

```

RESULT 11
PCT-US02-04432-7
; Sequence 7, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04432-7

```

```

Query Match          100.0%; Score 109; DB 1; Length 42;
Best Local Similarity 100.0%; Pred. No. 8e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RRVRRRRRRRRRRRRRRRRRRRRR 24
Db 1 RRVRRRRRRRRRRRRRRRRRRRRR 24

```

```

RESULT 12
PCT-US02-04812-7

```

```
RESULT 2
PCT-US02-04812-5
; Sequence 5, Application PC/TUS0204812
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04812
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04812-5

Query Match
Best Local Similarity 100.0%; Score 109; DB 1; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 3
US-09-785-058-5
; Sequence 5, Application US/09785058
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-5

Query Match
Best Local Similarity 100.0%; Score 109; DB 21; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 4
US-09-785-059-5
; Sequence 5, Application US/09785059
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-5

Query Match
Best Local Similarity 100.0%; Score 109; DB 21; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 5
US-10-079-075-5
; Sequence 5, Application US/10079075
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-5

Query Match
Best Local Similarity 100.0%; Score 109; DB 24; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 6
PCT-US02-04432-6
; Sequence 6, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04432-6

Query Match
Best Local Similarity 100.0%; Score 109; DB 1; Length 36;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 13 RRVRRRVRRRVRRRVRRRVRR 36

RESULT 7
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:55:47 ; Search time 103.915 Seconds
(without alignments)
148.906 Million cell updates/sec

Title: US-10-079-075-5
Perfect score: 109
Sequence: 1 RRVRRVRRVRRVRRVRRVRR 24

Scoring table: BL0SUM62
Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents AA Main:*

1:	/cgn2_6/prodata/1/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/prodata/1/paa/US06_COMB.pep.*
3:	/cgn2_6/prodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/prodata/1/paa/US08_COMB.pep.*
5:	/cgn2_6/prodata/1/paa/US081_COMB.pep.*
6:	/cgn2_6/prodata/1/paa/US082_COMB.pep.*
7:	/cgn2_6/prodata/1/paa/US083_COMB.pep.*
8:	/cgn2_6/prodata/1/paa/US084_COMB.pep.*
9:	/cgn2_6/prodata/1/paa/US085_COMB.pep.*
10:	/cgn2_6/prodata/1/paa/US086_COMB.pep.*
11:	/cgn2_6/prodata/1/paa/US087_COMB.pep.*
12:	/cgn2_6/prodata/1/paa/US088_COMB.pep.*
13:	/cgn2_6/prodata/1/paa/US090_COMB.pep.*
14:	/cgn2_6/prodata/1/paa/US091_COMB.pep.*
15:	/cgn2_6/prodata/1/paa/US092_COMB.pep.*
16:	/cgn2_6/prodata/1/paa/US093_COMB.pep.*
17:	/cgn2_6/prodata/1/paa/US094_COMB.pep.*
18:	/cgn2_6/prodata/1/paa/US095_COMB.pep.*
19:	/cgn2_6/prodata/1/paa/US096_COMB.pep.*
20:	/cgn2_6/prodata/1/paa/US097_COMB.pep.*
21:	/cgn2_6/prodata/1/paa/US098_COMB.pep.*
22:	/cgn2_6/prodata/1/paa/US099_COMB.pep.*
23:	/cgn2_6/prodata/1/paa/US100_COMB.pep.*
24:	/cgn2_6/prodata/1/paa/US101_COMB.pep.*
25:	/cgn2_6/prodata/1/paa/US102_COMB.pep.*
26:	/cgn2_6/prodata/1/paa/US103_COMB.pep.*
27:	/cgn2_6/prodata/1/paa/US104_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	24	1	PCT-US02-04432-5
2	109	100.0	24	1	PCT-US02-04812-5
3	109	100.0	24	21	US-09-785-058-5
4	109	100.0	24	21	US-09-785-059-5
5	109	100.0	24	24	US-10-079-075-5
6	109	100.0	36	1	PCT-US02-04432-6

7	109	100.0	36	1	PCT-US02-04812-6	Sequence 6, Appli
8	109	100.0	36	21	US-09-785-058-6	Sequence 6, Appli
9	109	100.0	36	21	US-09-785-059-6	Sequence 6, Appli
10	109	100.0	36	24	US-10-079-075-6	Sequence 6, Appli
11	109	100.0	42	1	PCT-US02-04432-7	Sequence 7, Appli
12	109	100.0	42	1	PCT-US02-04812-7	Sequence 7, Appli
13	109	100.0	42	21	US-09-785-058-7	Sequence 7, Appli
14	109	100.0	42	21	US-09-785-059-7	Sequence 7, Appli
15	109	100.0	42	24	US-10-079-075-7	Sequence 7, Appli
16	109	100.0	48	1	PCT-US02-04432-8	Sequence 8, Appli
17	109	100.0	48	1	PCT-US02-04812-8	Sequence 8, Appli
18	109	100.0	48	21	US-09-785-058-8	Sequence 8, Appli
19	109	100.0	48	21	US-09-785-059-8	Sequence 8, Appli
20	109	100.0	48	24	US-10-079-075-8	Sequence 8, Appli
21	88	80.7	24	1	PCT-US02-04432-10	Sequence 10, Appli
22	88	80.7	24	1	PCT-US02-04812-10	Sequence 10, Appli
23	88	80.7	24	21	US-09-785-058-10	Sequence 10, Appli
24	88	80.7	24	21	US-09-785-059-10	Sequence 10, Appli
25	88	80.7	24	24	US-10-079-075-10	Sequence 10, Appli
26	88	80.7	36	1	PCT-US02-04432-11	Sequence 11, Appli
27	88	80.7	36	1	PCT-US02-04812-11	Sequence 11, Appli
28	88	80.7	36	21	US-09-785-058-11	Sequence 11, Appli
29	88	80.7	36	21	US-09-785-059-11	Sequence 11, Appli
30	88	80.7	36	24	US-10-079-075-11	Sequence 11, Appli
31	88	80.7	48	1	PCT-US02-04432-12	Sequence 12, Appli
32	88	80.7	48	1	PCT-US02-04812-12	Sequence 12, Appli
33	88	80.7	48	21	US-09-785-058-12	Sequence 12, Appli
34	88	80.7	48	21	US-09-785-059-12	Sequence 12, Appli
35	88	80.7	48	24	US-10-079-075-12	Sequence 12, Appli
36	59	54.1	28	15	US-09-157-583-17	Sequence 17, Appli
37	57.5	52.8	882	18	US-09-416-195-78	Sequence 78, Appli
38	57.5	52.8	882	24	US-10-020-359B-94	Sequence 94, Appli
39	57	52.3	17	1	PCT-US96-03490-4	Sequence 4, Appli
40	57	52.3	17	8	US-08-415-239-4	Sequence 4, Appli
41	57	52.3	17	8	US-08-457-912-4	Sequence 4, Appli
42	57	52.3	17	13	US-08-968-929-4	Sequence 4, Appli
43	56	51.4	31	1	PCT-US02-04432-2	Sequence 2, Appli
44	56	51.4	31	1	PCT-US02-04812-2	Sequence 2, Appli
45	56	51.4	31	21	US-09-785-058-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
PCT-US02-04432-5
; Sequence 5, Application PC/TUS0204432
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mielzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-PCT / 072396.0223
; CURRENT APPLICATION NUMBER: PCT/US02/04432
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
PCT-US02-04432-5

Query Match 100.0%; Score 109; DB 1; Length 24;
Best Local Similarity: 100.0%; Pred. No. 4.2e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 RRVRRVRRVRRVRRVRRVRR 24
1 RRVRRVRRVRRVRRVRRVRR 24

```
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-8

Query Match      100.0%; Score 109; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.9e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 7 RRVRRRVRRRVRRRVRRRVRR 30

RESULT 13
US-09-785-059-10
; Sequence 10, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-10

Query Match      80.7%; Score 88; DB 9; Length 24;
Best Local Similarity 87.5%; Pred. No. 2.8e-06;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 14
US-10-079-075-10
; Sequence 10, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-10

Query Match      80.7%; Score 88; DB 9; Length 24;
Best Local Similarity 87.5%; Pred. No. 2.8e-06;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 15
US-09-785-058-10
; Sequence 10, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-10

Query Match      80.7%; Score 88; DB 9; Length 24;
Best Local Similarity 87.5%; Pred. No. 2.8e-06;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 1 RRVRRRVRRRVRRRVRRRVRR 24

Search completed: June 9, 2003, 12:34:09
Job time : 13.766 secs
```

FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 42
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-7

Query Match 100.0%; Score 109; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 8
US-10-079-075-7

Sequence 7, Application US/10079075
Publication No. US2002018102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 42
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-7

Query Match 100.0%; Score 109; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 9
US-09-785-058-7

Sequence 7, Application US/09785058
Publication No. US2003003627A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A 34001 / 072396.0222
CURRENT APPLICATION NUMBER: US/09/785.058
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 42
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-7

Query Match 100.0%; Score 109; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 10

US-09-785-059-8
Sequence 8, Application US/09785059
Patent No. US20020169279A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-8

Query Match 100.0%; Score 109; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.9e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30

RESULT 11

US-10-079-075-8
Sequence 8, Application US/10079075
Publication No. US2002018102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 48
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-8

Query Match 100.0%; Score 109; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 7.9e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRRVRRRVRRRVRRRVRR 24
DB 7 RRVRRRVRRRVRRRVRRRVRR 30

RESULT 12

US-09-785-058-8
Sequence 8, Application US/09785058
Publication No. US2003003627A1


```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-5

Query Match          100.0%; Score 109; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 3
; Sequence 5, Application US/09785058
; Publication No. US2003003627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-5

Query Match          100.0%; Score 109; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 1 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 4
; Sequence 6, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-6

Query Match          100.0%; Score 109; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Cy 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 13 RRVRRRVRRRVRRRVRRRVRR 36

RESULT 5
; Sequence 6, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-6

Query Match          100.0%; Score 109; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 13 RRVRRRVRRRVRRRVRRRVRR 36

RESULT 6
; Sequence 6, Application US/09785058
; Publication No. US2003003627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-6

Query Match          100.0%; Score 109; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRVRRRVRRRVRRRVRRRVRR 24
Db 13 RRVRRRVRRRVRRRVRRRVRR 36

RESULT 7
; Sequence 7, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 12:01:36 ; Search time 12.766 Seconds
(without alignments)
194.092 Million cell updates/sec

Title: US-10-079-075-5
Perfect score: 109
Sequence: 1 RRVRRVRRVRRVRRVRRVRR 24

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	24	US-09-785-059-5	Sequence 5, Appl1
2	109	100.0	24	US-10-079-075-5	Sequence 5, Appl1
3	109	100.0	24	US-09-785-058-5	Sequence 5, Appl1
4	109	100.0	36	US-09-785-059-6	Sequence 6, Appl1
5	109	100.0	36	US-10-079-075-6	Sequence 6, Appl1
6	109	100.0	36	US-09-785-058-6	Sequence 6, Appl1
7	109	100.0	42	US-09-785-059-7	Sequence 7, Appl1
8	109	100.0	42	US-10-079-075-7	Sequence 7, Appl1
9	109	100.0	42	US-09-785-058-7	Sequence 7, Appl1
10	109	100.0	48	US-09-785-059-8	Sequence 8, Appl1
11	109	100.0	48	US-10-079-075-8	Sequence 8, Appl1
12	109	100.0	48	US-09-785-058-8	Sequence 8, Appl1
13	109	100.0	24	US-09-785-059-10	Sequence 10, Appl1
14	88	80.7	24	US-10-079-075-10	Sequence 10, Appl1
15	88	80.7	24	US-09-785-058-10	Sequence 10, Appl1
16	88	80.7	36	US-09-785-059-11	Sequence 11, Appl1
17	88	80.7	36	US-10-079-075-11	Sequence 11, Appl1
18	88	80.7	36	US-09-785-058-11	Sequence 11, Appl1
19	88	80.7	48	US-09-785-059-12	Sequence 12, Appl1

20	88	80.7	48	9	US-10-079-075-12	Sequence 12, Appl1
21	88	80.7	48	9	US-09-785-058-12	Sequence 12, Appl1
22	56	51.4	31	9	US-09-785-059-2	Sequence 2, Appl1
23	56	51.4	31	9	US-10-079-075-2	Sequence 2, Appl1
24	56	51.4	31	9	US-09-785-058-2	Sequence 2, Appl1
25	55	50.5	28	9	US-09-785-059-1	Sequence 1, Appl1
26	55	50.5	28	9	US-10-079-075-1	Sequence 1, Appl1
27	55	50.5	28	9	US-09-785-058-1	Sequence 1, Appl1
28	54	49.5	12	9	US-09-785-059-4	Sequence 4, Appl1
29	54	49.5	12	9	US-10-079-075-4	Sequence 4, Appl1
30	54	49.5	12	9	US-09-785-058-4	Sequence 4, Appl1
31	49	45.0	18	10	US-09-840-009-36	Sequence 36, Appl1
32	49	45.0	18	10	US-09-840-009-37	Sequence 37, Appl1
33	47	43.1	12	9	US-09-785-059-9	Sequence 9, Appl1
34	47	43.1	12	9	US-10-079-075-9	Sequence 9, Appl1
35	47	43.1	12	9	US-09-785-058-9	Sequence 9, Appl1
36	46	42.2	70	10	US-09-764-877-1553	Sequence 1553, Appl1
37	46	42.2	997	9	US-10-176-847-50	Sequence 50, Appl1
38	46	42.2	999	10	US-09-747-371-2	Sequence 2, Appl1
39	45	41.3	31	9	US-09-785-059-3	Sequence 3, Appl1
40	45	41.3	31	9	US-10-079-075-3	Sequence 3, Appl1
41	45	41.3	31	9	US-09-785-058-3	Sequence 3, Appl1
42	44	40.4	18	9	US-10-060-102-10	Sequence 10, Appl1
43	44	40.4	18	9	US-10-060-102-11	Sequence 11, Appl1
44	44	40.4	29	9	US-10-060-102-8	Sequence 8, Appl1
45	43.5	39.9	96	9	US-10-093-892-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-09-785-059-5
; Sequence 5, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; SOFTWARE: FASTSEQ for Windows Version 3.0
; NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-5

Query Match 100.0%; Score 109; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRRVRRVRRVRR 24
DB 1 RRVRRVRRVRRVRRVRRVRR 24

RESULT 2
US-10-079-075-5
; Sequence 5, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12

CC The present sequence is one of 169 disclosed specific examples of
CC the new peptides. It is an analogue of the peptide designated L2P1
CC (see AA47614) which is a peptide from the transmembrane protein (gp41)
CC of HIV strain HXB2R.

SQ **Sequence** **21 AA;**

Query Match	45.0%	Score 49;	DB 19;	Length 21;
Best Local Similarity	47.1%	Pred. No. 2.3;		
Matches	8;	Conservative	6;	Mismatches 3.;
			Indels	0;
			Gaps	0

Qy 6 RVRVRVRVRVRV 22

Db 1 RIRRPRIICIGQV 17

RESULT 15

ID	AYY32598 standard; peptide; 21 AA.
xx	

AC AAY32598;
VY

DT 21-OCT-1999 (first entry)
 VV

Antimicrobial peptide LLP1 analogue

KW Antimicrobial peptide; LBP1; SLP-1; LBP2; SLP2A; SLP2B; ELP; infection;
KW growth inhibitor; microorganism; virus; gene therapy; vector production;
KW sterilisation.

05 Synthetic.
06

XX
XX
TTCG04EE67-1
DNXX 31-ATTC-1000
DDXX
DE 18 SEP 1967

XX
DB 36 JAN 1965 0610634

PR 24-JAN-1997; 97US-0786748.

XX
(IVDT) / IVTII STECHNICHU

	XX	PT	Mt + max	TB	Montrol zero
0					
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

XX WPT 1000-500100/43
DB

XX
XX
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
8

XX Disclosures: Column 1

Feb 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

CC is an analogue of the peptide LLPI (N

melittin resistant *S. aureus*, *Pseudomonas aeruginosa*, *Enterococcus faecalis*, *S. marcescens*, *Escherichia coli*, fungi, protozoa and viruses in a mammalian host. They can be used to inhibit growth of diverse microorganisms such as bacteria, fungi, protozoa and DNA and RNA viruses and can be used in tissue culture to inhibit unwanted microbial growth, particularly for the production of recombinant proteins or vectors for gene therapy. They can also be used in preventing infections through the sterilisation of wounds prior to suture and to sterilise surgical instruments. The unique structure of these antimicrobial peptides imparts high potency while selectivity is maintained, they are moderately haemolytic but only lyse red blood cells at high concentrations unlike melittin, a peptide extracted from bee venom, which is highly active against bacteria and lyses red blood cells showing little selectivity. The peptides target a membrane structure which makes it more difficult for a microorganism to develop a mechanism of resistance against this type of antibiotic. Their small size makes them relatively simple to prepare by standard synthetic peptide chemistry.

CC emergence of resistance to other inhaled antibiotics; instillation into
 CC the urinary bladder of patients with indwelling catheters to prevent
 CC infection; application to the skin of patients with serious burns;
 CC ophthalmic instillation, directly or in ophthalmic solutions, to treat or
 CC prevent infection; and intravaginal application to treat bacterial
 CC vaginosis and/or prevent sexually transmitted disease, e.g. by preventing
 CC infection with *Chlamydia trachomatis*. The novispirins also find use in
 CC the treatment of plant-pathogenic pseudomonads, in agricultural
 CC applications designed to prevent disease in and spoilage of food crops.
 CC Novispirin is useful in vitro formulations to kill microbes, where the
 CC use of conventional antibiotics is not desirable, e.g., novispirins may
 CC be added to animal and/or human food preparations, and as an additive for
 CC in vitro cultures of cells, to prevent the overgrowth of microbes in
 CC tissue culture. Novispirin is also useful for killing non-bacterial
 CC pathogens such as fungal and protozoan pathogens. The invention is useful
 CC for treating a host suffering from or predisposed to a microbial
 CC infection. The peptides are nonhaemolytic, exhibit reduced in vitro
 CC cytotoxicity relative to other antimicrobial peptides and are well-
 CC tolerated in vivo after intravenous injection. Novispirins are equally
 CC effective against growing and stationary phase *Pseudomonas aeruginosa* and
 CC they retain activity in the presence of high concentrations of salt or
 CC human serum. Novispirins also bind lipopolysaccharide (LPS), a property
 CC that may mitigate symptoms associated with gram-negative bacterial
 CC infection. The present sequence is R1, R2, G10-novispirin amide peptide.

XX SQ Sequence 18 AA;
 Query Match 45.0%; Score 49; DB 23; Length 18;
 Best Local Similarity 37.5%; Pred. No. 2;
 Matches 6; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

OY 5 RRVRRVRRVRRVRR 20
 DB 1 RRRRRIRKGRIRIKK 16

RESULT 13
 AAM47663 standard; peptide; 21 AA.

XX AC AAM47663;
 XX 26-MAY-1998 (first entry)
 DE Antimicrobial peptide ILPI analogue.
 XX
 KM Antimicrobial; transmembrane protein; TM; lentivirus lytic peptide;
 KM LIP; amphipathic; antibacterial; antifungal; antiviral; antiprotozoal.
 OS Synthetic.
 OS Human immunodeficiency virus.
 XX PN US5714577-A.
 XX PD 03-FEB-1998.
 XX PF 24-JAN-1997; 97US-0786748.
 XX PR 26-JAN-1996; 96US-0010634.
 XX PR 24-JAN-1997; 97US-0786748.
 XX PA (UYPI-) UNIV PITTSBURGH.
 XX PI Mietzner TA, Montelaro RC, Tencza SB;
 XX DR WPI; 1998-158352/14.
 XX PT Retroviral TM peptides - useful as antibacterial agents.
 XX PS Disclosure; Column 9; 59pp; English.
 CC The invention relates to new antimicrobial peptides which correspond to
 CC amino acid sequences in the transmembrane proteins of lentiviruses, in

CC particular HIV and SIV. These peptides comprise arginine rich sequences
 CC which, when modelled for secondary structure, display high
 CC amphipathicity and hydrophobic moment. Also disclosed are structural
 CC and functional analogues and homologues of these peptides which also
 CC display antimicrobial activity. The peptides are highly inhibitory to
 CC microorganisms (bacteria, fungi, viruses and protozoa) but significantly
 CC less toxic to red blood cells and other normal mammalian cells. Activity
 CC is demonstrated against Gram positive and negative bacteria including
 CC *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Enterococcus faecalis* and
 CC *Serratia marcescens*.

XX SQ Sequence 21 AA;
 Query Match 45.0%; Score 49; DB 19; Length 21;
 Best Local Similarity 44.4%; Pred. No. 2.3;
 Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 6 RRVRRVRRVRRVRRVRR 23
 DB 1 RRRRRIRKGRIRIKK 18

RESULT 14
 AAM47664 standard; peptide; 21 AA.

XX AC AAM47664;
 XX 26-MAY-1998 (first entry)
 DE Antimicrobial peptide ILPI analogue.
 XX
 KM Antimicrobial; transmembrane protein; TM; lentivirus lytic peptide;
 KM LIP; amphipathic; antibacterial; antifungal; antiviral; antiprotozoal.
 OS Synthetic.
 OS Human immunodeficiency virus.
 XX PN US5714577-A.
 XX PD 03-FEB-1998.
 XX PF 24-JAN-1997; 97US-0786748.
 XX PR 26-JAN-1996; 96US-0010634.
 XX PR 24-JAN-1997; 97US-0786748.
 XX PA (UYPI-) UNIV PITTSBURGH.
 XX PI Mietzner TA, Montelaro RC, Tencza SB;
 XX DR WPI; 1998-158352/14.
 XX PT Retroviral TM peptides - useful as antibacterial agents
 XX PS Disclosure; Column 9; 59pp; English.
 CC The invention relates to new antimicrobial peptides which correspond to
 CC amino acid sequences in the transmembrane proteins of lentiviruses, in
 CC particular HIV and SIV. These peptides comprise arginine rich sequences
 CC which, when modelled for secondary structure, display high
 CC amphipathicity and hydrophobic moment. Also disclosed are structural
 CC and functional analogues and homologues of these peptides which also
 CC display antimicrobial activity. The peptides are highly inhibitory to
 CC microorganisms (bacteria, fungi, viruses and protozoa) but significantly
 CC less toxic to red blood cells and other normal mammalian cells. Activity
 CC is demonstrated against Gram positive and negative bacteria including
 CC *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Enterococcus faecalis* and
 CC *Serratia marcescens*.

RESULT 6

AAR92435

ID AAR92435 standard; peptide; 27 AA.

AC AAR92435;

DT 18-SEP-1996 (first entry)

DE Lytic peptide used in ubiquitin-lytic peptide fusion protein.

KM Ubiquitin; fusion protein; lysis; infection; neoplasia; wound healing;

KW stability; reduced toxicity.

OS Synthetic.

PN WO9603519-A1.

PD 08-FEB-1996.

PF 24-JUL-1995; 95WO-US09339.

PR 22-JUL-1994; 94US-0279472.

PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.

XX (USDA) US SEC OF AGRIC.

PI Belknap W, Garbano J, Jaynes J;

DR WPI; 1996-117061/12.

PT New fusion protein of ubiquitin and a lytic peptide - for treating

PT infections and neoplasia, heating wounds, etc. also related nucleic

PS acid, vectors, and transformed cells

XX Claim 5; Page 26; 112pp; English.

CC AAR93372-R92462 are lytic peptides used to create ubiquitin-lytic

CC peptide fusion proteins in which the ubiquitin polypeptide is linked

CC at its 3'-terminus to the lytic peptide. The lytic peptides are pref-

CC selected from either the cecropins, defensins, sarcotoxins, melittin

CC and magainins. The fusion proteins (FPs) are useful for treating

CC protozoal, bacterial, fungal and viral infections and neoplasia (in

CC plants and animals) in the same way as the FP alone, they also

CC promote wound healing. FPs produced in bacteria may be cleaved in

CC vitro by ubiquitin hydrolases to recover the active lytic peptide.

CC FPs produced in eukaryotic cells are cleaved by endogenous enzymes

CC to yield lytic peptide. Recombinant DNA encoding the FPs have.

CC greater stability in bacteria than DNA encoding the lytic peptide

CC only.

SQ Sequence 27 AA;

Query Match 47.7%; Score 52; DB 17; Length 27;

Best Local Similarity 25.0%; Pred. No. 1.2; Matches 6; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRRVRRVRRVRR 24

DB 3 RRIKKRIKKIVRRFRIALIRK 26

RESULT 7

AAR89992

ID AAR89992 standard; peptide; 27 AA.

AC AAR89992;

DT 16-SEP-1996 (first entry)

DE Synthetic lytic peptide #26.

KW Lytic peptide; ubiquitin; synthetic analogue; cell membrane; cell lysis;

KW microbial pathogen; disease-resistant plant; bacterial infection; fungus;

KW protozoa; virus; neoplasia; fusion protein; hydrolase.

OS Synthetic.

PN WO9603522-A1.

PD 08-FEB-1996.

PF 24-JUL-1995; 95WO-US09338.

PR 22-JUL-1994; 94US-0279472.

PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.

XX (DEME-) DEMETER BIOTECHNOLOGIES LTD.

PI Jaynes J;

DR WPI; 1996-117064/12.

PT Lytic peptide(s), useful for developing disease-resistant plants -

PT can be expressed as fusion protein with ubiquitin for stable prodn.

PS Claim 1; Page 81; 111pp; English.

CC AAR89967-R90021 and AAR90726-R90763 represent synthetic analogues of

CC naturally occurring lytic peptides. Lytic peptides destroy prokaryotic

CC and other non-host cells by disrupting the cell membrane and promoting

CC cell lysis. Synthetic lytic peptide analogues have similar or higher

CC levels of lytic activity for many different types of cells, compared to

CC naturally occurring forms. The concentration of the synthetic analogue

CC required to lyse microbial pathogens does not lyse normal mammalian

CC cells. The lytic peptides can be expressed in plants to allow for the

CC development of disease-resistant plants. The peptides are useful in

CC promoting wound healing and combating bacterial infections in plants.

CC The lytic peptides can also be used for combating protozoal, fungal,

CC viral or bacterial infections or neoplasias in mammals and plants.

CC Lytic peptide-ubiquitin fusion proteins are suitable for production in

CC bacterial hosts. Bacteria lack the hydrolase which cleaves the peptide

CC from ubiquitin, and therefore the active (and cytotoxic) lytic peptide

CC will not be released in the host cells. The recombinantly produced

CC lytic peptide can be retrieved from the fusion protein by cleavage in

CC vitro.

SQ Sequence 27 AA;

Query Match 47.7%; Score 52; DB 17; Length 27;

Best Local Similarity 25.0%; Pred. No. 1.2; Matches 6; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRRVRRVRRVRR 24

DB 3 RRIKKRIKKIVRRFRIALIRK 26

RESULT 8

AA330017

ID AA330017 standard; peptide; 21 AA.

AC AA330017;

DT 29-SEP-1999 (first entry)

DE Bufoin II-derived antimicrobial peptide.

KM Bufoin II; antimicrobial; alpha-helix; salt insensitive; magainin;

KW gram-negative bacteria; gram-positive bacteria; fungi; protozoa.

OS Synthetic.

PN WO9337664-A1.

XX Example 102; Page 84; 119pp; English.

CC The invention provides an antimicrobial compound (I) which is a peptide
XX having 8-50 amino acids, a net charge of 4, a hydrophobic moment (micron)
CC as a beta sheet which is 0.2 higher than its micro H as alpha helix, and
CC having detectable membrane disrupting activity against a microbial
CC pathogen, and substantially no membrane disrupting activity against
CC mammalian cells. (I) is useful for inhibiting microbial activity. (I)
CC has a detectable membrane disrupting activity against a pathogen, and is
CC useful for killing human sperm. The peptides are also provided in the
CC form of an expression vector comprising a nucleic acid encoding the
CC peptide. The peptides are useful for inhibiting the activity of
CC bacteria, and other microbial pathogens such as algae, fungi or protozoa
CC and for inhibiting non-microbial pathogens such as worms or arthropods,
CC and as spermicides for humans as the sperm membrane is atypical of human
CC cell membranes. (I) also has diagnostic uses e.g., in localizing an
CC infection or detecting sepsis. The peptides may act as binding molecules
CC and are useful to purify a target from blood, for qualitative or
CC quantitative analysis of analytes in in vitro sample, and for in vivo
CC imaging. Also, they are useful as molecular weight markers, as nutrient
CC source, as growth medium component for culturing microorganisms, as well
CC as a food ingredient for human consumption. The peptides have a greater
CC selectivity for bacterial versus mammalian lipids as compared to the
CC alpha helical peptides. Sequences AAG5536-47 represent amino acid
CC sequences of antimicrobial peptides.

XX Sequence 18 AA;

Query Match 58.7%; Score 64; DB 22; Length 18;
Best Local Similarity 50.0%; Pred. No. 0.022;
Matches 8; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

QY 9 RVRVRVRRVRRVRR 24
|::||:|::||:|::||
Db 2 RIRRRIRIRIRIR 17

RESULT 2
AAP91336
ID AAP91336 standard; peptide; 28 AA.
XX
AC AAP91336;
XX
DT 19-MAR-1990 (first entry)
XX
DE Amino acid sequence of Shiva-4.
XX
KM Shiva-4; lytic peptide; antimicrobial peptide; disease-resistant
KM trichophyte; Shiva-2; Shiva-3; Shiva-5; Shiva-6; Shiva-7.
XX
PN WO8904371-A.
XX
PD 18-MAY-1989.
XX
PF 02-NOV-1988; 88WO-US03908.
XX
PR 02-NOV-1987; 87US-0115941.
XX
PA (LOU) LOUISIANA STATE UNIV.
XX
PI Jaynes JM, Derrick KS;
XX
DR WPI; 1989-165650/22.
XX
PT Transformed plants contg. heterologous gene - expressing antimicrobial
CC agent, or polypeptide high in essential amino acids
XX
PS Tiple I; 56pp; English.
XX
CC Amino acid sequence of Shiva-4 as an exemplary lytic peptide for
CC use as an antimicrobial peptide contemplated for use in plant

CC (trichophyte) transformants in the invention. It is a homologue of
CC Shiva-2, -3 and -5 to -7. All of these Shiva peptides are also
CC contemplated as having general utility in inducing lysis of cells in
CC vitro. Shiva-4 may be too lytically active to be used in plants at high
CC expression levels.

XX Sequence 28 AA;

Query Match 54.1%; Score 59; DB 10; Length 28;
Best Local Similarity 45.5%; Pred. No. 0.15;
Matches 10; Conservative 11; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRVRVRRVRRVRRVRR 23
|::||:|::||:|::||
Db 6 RLRLRLRLRLRLRLRL 27

RESULT 3
AAW05116
ID AAW05116 standard; peptide; 17 AA.
XX
AC AAW05116;
XX
DT 19-JUN-1997 (first entry)
XX
DE Porcine somatotropin mimic.
XX
KM porcine somatotropin; pST; growth promoter; helical conformation.
XX
OS Synthetic.
XX
PN WO9630405-A1.
XX
PD 03-OCT-1996.
XX
PF 15-MAR-1996; 96WO-US03490.
XX
PR 31-MAR-1995; 95US-0415239.
XX
PA (AMCY) AMERICAN CYANAMID CO.
XX
PI Buckwalter BL, Shieh H, Wang BS;
XX
DR WPI; 1996-485447/48.
XX
PT Peptide(s) mimicking a helical region of porcine somatotropin - used
PT in compositions to promote mammalian growth
XX
PS Claim 3; Page 17; 63pp; English.
XX
CC New peptides are disclosed which, by virtue of having certain
CC defined amino acids at every third or fourth residue, have a well
CC defined secondary structure which mimics the helical conformation
CC of a corresponding region of porcine somatotropin (pST). The peptides
CC enhance the activity of pST and promote the growth of warm-blooded
CC animals, especially pigs. They compete with pST for binding to the
CC BS-7.6 monoclonal antibody. The peptides have the generic sequence
CC XXXIXXXXXXXVXX (I) or XXXIXXXXXXXIXXXV (II); where residues X are
CC undefined other than the statement that the sequences differ from the
CC native sequence of pST. Formula (II) represents a peptide in which the
CC location of the essential amino acids is shifted by three amino acids,
CC representing almost one turn along the helix. Preferably X(2) of (II) is
CC Ile. Preferably the peptides contain Ser (as a promoter of helical
CC conformation) as the amino acid immediately amino-terminal to the first
CC Leu in (I) or to the first Ile of (II). Also, one or more of the first
CC or second Leu or the Val of (I) may be replaced by Nle. Furthermore, a
CC Cys residue may be added to either or both ends of the peptides.
CC The present sequence represents a specific example of the new
CC peptides.

XX Sequence 17 AA;

Query Match 52.3%; Score 57; DB 17; Length 17;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:35:06 ; Search time 25.8723 Seconds
(without alignments)
123.607 Million cell updates/sec

Title: US-10-079-075-5
Perfect score: 109
Sequence: 1 RRVRRVRRVRRVRRVRRVRR 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	58.7	18	22	AA65539
2	59	54.1	28	10	AA691336
3	57	52.3	17	17	AAW05116
4	55	50.5	40	16	AAAR4926
5	54	49.5	39	18	AAW06684
6	52	47.7	27	17	AAAR92435
7	52	47.7	27	17	AAAR9992
8	51	46.8	21	20	AAAY30017
9	51	46.8	31	15	AAAR60065
10	51	46.8	31	16	AAAR80735

11	49	45.0	18	23	AAE20888	R1, R2, G10-novisp
12	49	45.0	18	23	AAE20889	R1, R2, G10-novisp
13	49	45.0	21	19	AAW47663	Antimicrobial pept
14	49	45.0	21	19	AAW47664	Antimicrobial pept
15	49	45.0	21	20	AAAY32598	Antimicrobial pept
16	49	45.0	21	20	AAAY32599	Antimicrobial pept
17	49	45.0	23	17	AAAR92434	Antimicrobial pept
18	49	45.0	23	17	AAAR9991	Antimicrobial pept
19	49	45.0	24	19	AAW47666	Antimicrobial pept
20	49	45.0	24	20	AAAY32601	Antimicrobial pept
21	49	45.0	27	17	AAAR92436	Antimicrobial pept
22	49	45.0	27	17	AAAR9993	Antimicrobial pept
23	47	43.1	24	19	AAW47669	Antimicrobial pept
24	47	43.1	24	19	AAW47671	Antimicrobial pept
25	47	43.1	24	20	AAAY32604	Antimicrobial pept
26	47	43.1	24	20	AAAY32606	Antimicrobial pept
27	47	43.1	28	19	AAW47652	Antimicrobial pept
28	47	43.1	28	19	AAW47655	Antimicrobial pept
29	47	43.1	28	19	AAW47657	Antimicrobial pept
30	47	43.1	28	20	AAAY32592	Antimicrobial pept
31	47	43.1	28	20	AAAY32597	Antimicrobial pept
32	47	43.1	28	20	AAAY32590	Antimicrobial pept
33	47	43.1	516	20	AAV07735	Human breast-speci
34	46	42.2	21	19	AAW47662	Antimicrobial pept
35	46	42.2	21	19	AAW47659	Antimicrobial pept
36	46	42.2	21	20	AAAY32594	Antimicrobial pept
37	46	42.2	21	20	AAAY32597	Antimicrobial pept
38	46	42.2	24	19	AAW47667	Antimicrobial pept
39	46	42.2	24	20	AAAY32602	Antimicrobial pept
40	46	42.2	27	16	AAAR84148	Antimicrobial pept
41	46	42.2	27	16	AAAR77062	Antimicrobial pept
42	46	42.2	27	16	AAAR74711	Antimicrobial pept
43	46	42.2	27	16	AAAR64790	Antimicrobial pept
44	46	42.2	27	17	AAAR92392	Antimicrobial pept
45	46	42.2	27	17	AAAR90746	Antimicrobial pept

ALIGNMENTS

RESULT 1

AA65539
ID AA65539 standard; peptide, 18 AA.

AC AA65539;
XX

DT 30-NOV-2001 (first entry)
XX

DE Peptide sequence used in the course of the invention.
XX

KW Antimicrobial; microbial membrane disrupter; gene therapy; pathogen;
XX

KW spermidine; imaging; magainin; PGla.
XX

OS Synthetic.
XX

PN WO200160162-A2.
XX

PD 23-AUG-2001.
XX

PF 15-FEB-2001; 2001WO-US04822.
XX

PR 15-FEB-2000; 2000US-0182495.
XX

PA (UYOH-) UNIV OHIO.
XX

PI Biazylk JF;
XX

DR WPI; 2001-565322/63.
XX

PT Novel peptides having antimicrobial activity have positive charge to
selectively disrupt microbial membranes, assume beta sheet structure in
membrane environment and are substantially amphipathic in beta sheet
structure -

RC STRAIN=CV. NIPPONBARE;
 RA Ito Y.;
 RT "Oryza sativa dihydroorotate dehydrogenase gene, complete cds."
 RL Submitted (FE8-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB039557, BAB11988.1; -.
 DR InterPro: IPR001295; DHO_dh.
 DR InterPro: IPR003009; FMN_enzyme.
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF01180; DHODEHase; 1.
 DR TIGRfam: TIGR01036; pyrd_sub2; 1.
 DR PROSITE: PS00911; DHODEHASE_1; 1.
 DR PROSITE: PS00912; DHODEHASE_2; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 SQ SEQUENCE 468 AA; 50346 MW; DDE7C308DC4C71F3 CRC64;

Query Match 42.7%; Score 46.5; DB 10; Length 468;
 Best Local Similarity 46.9%; Pred. No. 1.3e+02;
 Matches 15; Conservative 2; Mismatches 4; Indels 11; Gaps 2;

Qy 2 RVRVRVRVRVRVRVRV-----RRVVR 24
 Db 23 RVARRARRQLRR--RAVGPARPPPKRRLVVR 52

Search completed: June 9, 2003, 12:01:05
 Job time : 24.0426 secs

RA Naife D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malynch A.G., Koonin E.V., Kozlovskiy S.A.,
 RT "The complete genome of Hyperthermophilic Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL: AE010369; AA001984.1; .
 KM Complete proteome.
 SQ SEQUENCE 331 AA; 36904 MW; 3B27C993839FA067 CRC64;

Query Match 43.1%; Score 47; DB 17; Length 331;
 Best Local Similarity 55.0%; Pred. No. 82;
 Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 5 RRVRRVRRVRRVRRVRR 24
 Db 175 RYRVRRVRRVRRVRRVRR 194

RESULT 12
 Q99AR5 PRELIMINARY; PRT; 760 AA.

AC Q99AR5; 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Orf1.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OC NCBI_Taxid=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TCN-G1;
 RA Luo K.-X., He H.-T., Liu D.-X., Liu Z.-H., Xiao H., Jiang X.-J.,
 RA Liang W.-F., Zhang L.;
 RT "Novel variants related to TT virus wide distribution in China."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF345521; AKL1696.1; .
 DR InterPro: IPR004219; TTVirus_unk.
 DR Pfam: PF02956; TT_ORF1.1.
 SQ SEQUENCE 760 AA; 90494 MW; 50BD115CCF5181A CRC64;

Query Match 43.1%; Score 47; DB 12; Length 760;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 12; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
 QY 1 RRVRRVRRVRRVRRVRR 24
 Db 25 RLPTRRVRRVRRVRRVRR 48

RESULT 13
 Q41556 PRELIMINARY; PRT; 841 AA.
 AC Q41556; 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C17;
 RX MEDLINE=98105804; PubMed=9445059;
 RA Connor R.I., Korber B.T.M., Graham B.S., Hahn B.H., Ho D.D.,
 RA Walker B.D., Neumann A.U., Vermond S.H., Mestecky J., Jackson S.,
 RA Fenamore E., Cao Y., Gao Y., Kaysams S., Kunstan K.D., McDonald D.,
 RA McWilliams N., Trkola A., Moore J.P., Molinsky S.M.;
 RT "Immunological and virological analyses of persons infected by human
 RT immunodeficiency virus type 1 while participating in trials of
 RT recombinant gp120 subunit vaccines.";

RL J. Virol. 72:1552-1576(1998).
 DR EMBL: U84814; AAC5884.1; .
 DR InterPro: IPR000328; Env_GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KM AIDS: Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 FT NON_TER
 SQ SEQUENCE 841 AA; 95802 MW; AAF104893D91667 CRC64;

Query Match 43.1%; Score 47; DB 15; Length 841;
 Best Local Similarity 45.0%; Pred. No. 2e+02;
 Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 QY 2 RRVRRVRRVRRVRRVRR 21
 Db 813 RVIEVLRVRRVRRVRRVRR 832

RESULT 14
 Q90CG7 PRELIMINARY; PRT; 862 AA.

AC Q90CG7; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Envelope glycoprotein.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMS3379;
 RX MEDLINE=21342588; PubMed=11448170;
 RA Carr J.K., Torimiro J.N., Wolfe N.D., Eitel M.N., Kim B.,
 RA Sanders-Buell E., Jagodzinski L.L., Gotte D., Burke D.S., Birx D.L.,
 RA McCutchan F.B.;
 RT "The AG recombinant IBNG and novel strains of group M HIV-1 are common
 RT in Cameroon."
 RL Virology 286:168-181(2001).
 DR EMBL: AF377859; AAK59217.1; .
 DR InterPro: IPR000328; Env_GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KM AIDS: Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 862 AA; 97328 MW; C93A1D272BB09961 CRC64;

Query Match 43.1%; Score 47; DB 15; Length 862;
 Best Local Similarity 47.4%; Pred. No. 2.1e+02;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 6 RRVRRVRRVRRVRRVRR 24
 Db 834 RVIEVLRVRRVRRVRRVRR 852

RESULT 15
 Q9FZM9 PRELIMINARY; PRT; 468 AA.
 AC Q9FZM9; 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Dihydroorotate dehydrogenase.
 OR Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriocaridaceae; Oryzaeae; Oryza.
 OC NCBI_Taxid=4530;
 RN [1]
 RP SEQUENCE FROM N.A.

RT no. 1 TTV genotypes in Indonesia.";
RL Arch. Virol. 146:1249-1266(2001).
DR EMBL; AB054648; BAB61611.1; -.
DR InterPro; IPR004219; TTVirus_unk.
DR Pfam; PF02956; TTV_ORF1.1.
SQ SEQUENCE 742 AA; 88114 MW; A4E27AB09163DB5A CRC64;

Query Match 45.9%; Score 50; DB 12; Length 742;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 15; Conservative 2; Mismatches 7; Indels 6; Gaps 1;

Qy 1 RRVRRVRR-----VVRVRRVRRVRR 24
Db 39 RRRVRLRRRRRRGMAARRRLRRRVRR 68

RESULT 5
Q9SHX2 PRELIMINARY; PRT; 349 AA.
AC Q9SHX2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE At2G06420 protein (Hypothetical 41.9 kDa protein).
GN At2G06420.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Roundley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.U., Roming C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanhaken S.E., Umeyam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Frazer C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RA [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanhaken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Roming C.M., Benito M.-I.,
RA Carreira A.J., Creasy T.H., Buell C.R., Town C.D., Niernan W.C.,
RA Frazer C.M., Venter J.C.;
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC007235; AAD26890.1; -.
DR EMBL; AC006918; AAM15311.1; -.
KM Hypothetical protein.
SQ SEQUENCE 349 AA; 41935 MW; DCE334C856C9F0F5 CRC64;

Query Match 45.4%; Score 49.5; DB 10; Length 349;
Best Local Similarity 48.0%; Pred. No. 41;
Matches 12; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Qy 1 RRVRRVRR-VVRVRRVRRVRR 24
Db 278 RRVRRVRRVRRVRRVRRVRR 302

RESULT 6
Q8TW69

ID Q8TW69 PRELIMINARY; PRT; 241 AA.
AC Q8TW69;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Uncharacterized protein.
GN MK1167.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatunov R.L., Wolf Y.I., Stetter K.O.,
RA Malynch A.G., Koonin E.V., Kozlyavkin S.A.,
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010407; AAM02380.1; -.
KM Complete proteome.
SQ SEQUENCE 241 AA; 25177 MW; 3A8ACFE5874735E1 CRC64;

Query Match 44.0%; Score 48; DB 17; Length 241;
Best Local Similarity 47.1%; Pred. No. 45;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 8 RRVRRVRRVRRVRR 24
Db 182 RETIRTVRLRRVLR 198

RESULT 7
Q987V7 PRELIMINARY; PRT; 349 AA.
AC Q987V7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein ml16891.
GN ML16891.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003010; BAB53093.1; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 349 AA; 37473 MW; B7E34BCECC39304 CRC64;

Query Match 44.0%; Score 48; DB 16; Length 349;
Best Local Similarity 68.8%; Pred. No. 64;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 RRVRRVRRVRRVRR 20
Db 106 RRIARGVRRVRRVRR 121

RC STRAIN=M8R-1;
RX MEDLINE=99119503; PubMed=9918888;
RA Reddy P.S., Chen Y., Idamakanti N., Pyne C., Babiuk L.A., Tikoo S.K.;
RT "Characterization of early region 1 and p1x of bovine adenovirus-3";
RL Virology 253:299-308 (1999).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M8R-1;
RA Reddy P.S., Idamakanti N., Zakhartchouk A.N., Baxi M.K., Lee J.B.,
RA Pyne C., Babiuk L.A., Tikoo S.K.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030154; AAD09724.1;
DR InterPro; IPR004912; Adeno VII;
DR Pfam; PF03228; Adeno VII; 1.
DR SEQUENCE 171 AA; 18959 MW; 0EACB1C3C12519A CRC64;
Query Match 49.1%; Score 53.5; DB 12; Length 171;
Best Local Similarity 51.6%; Pred. No. 6.2;
Matches 16; Conservative 3; Mismatches 5; Indels 7; Gaps 1;
Db 1 RRVRRVRRV-----VRRVRRVRRVRR 24
88 RRGVRRVRLRRSPRTLRRSVRQVRR 118
RESULT 2
08TXS5
ID 08TXS5 PRELIMINARY; PRT; 428 AA.
AC 08TXS5;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE Uncharacterized protein.
GN MK0585.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
ON NCBI_TaxID=2320;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Sheshberdana O.V., Shakhova V.V., Belova G.I., Arsvind L.,
RA Natsale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malynk A.G., Koonin E.V., Koznyavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).
DR EMBL; AB010352; AAM01800.1; -.
KW Complete proteome.
SQ SEQUENCE 428 AA; 48744 MW; 964D477CA264D13C CRC64;
Query Match 47.2%; Score 51.5; DB 17; Length 428;
Best Local Similarity 53.8%; Pred. No. 27;
Matches 14; Conservative 6; Mismatches 3; Indels 3; Gaps 2;
Db 1 RRV-RRV--RRVRRVRRVRRVRR 23
360 REVERRRVGRRLVRKIVRKIVRLX 385
RESULT 3
09PFU7
ID 09PFU7 PRELIMINARY; PRT; 240 AA.
AC 09PFU7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE GMA¹ synthase.
GN XF0560.
OS Xylella fastidiosa.
OC Bacteriay Proteobacteria; gamma subdivision; Xanthomonas group;

OC Xylella.
ON NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares A.R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barrios M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britons M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,
RA Colauto N.B., Colombo C., Costa F.P., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facchini A.P., Ferreira A.J.S., Franco M.C., Frohme M., Furlan L.R.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garner J., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.B., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pequeiro J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Trufi D., Tsai S.M., Tsunako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Vago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159 (2000).
DR EMBL; AE003803; AAP83370.1; -.
DR InterPro; IPR000991; GATase_1.
DR Pfam; PF00117; GATase; 1.
DR PRINTS; PR00096; GATase.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
KW Complete proteome.
SQ SEQUENCE 240 AA; 26350 MW; FF815EE1EE8A35 CRC64;
Query Match 45.9%; Score 50; DB 16; Length 240;
Best Local Similarity 46.4%; Pred. No. 24;
Matches 13; Conservative 4; Mismatches 5; Indels 6; Gaps 1;
Db 1 RRVRRV-----RRVRRVRRVRRV 22
213 RRIARQVSAAPVAVQVLRFRVRARRAV 240
RESULT 4
09IPB3
ID 09IPB3 PRELIMINARY; PRT; 742 AA.
AC 09IPB3;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ORF1.
OS TT virus.
OS Viruses; ssDNA viruses; unclassified ssDNA viruses.
ON NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT-10F;
RA Okamoto H.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21440417; PubMed=11556704;
RX Muijono D.H., Nishizawa T., Tsuda F., Takahashi M., Okamoto H.;
RT "Molecular epidemiology of TT virus (TTV) and characterization of two

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:43:01 ; Search time 22.0426 Seconds
(without alignments)
224.345 Million cell updates/sec

Title: US-10-079-075-5

Perfect score: 109
Sequence: 1 RRVRRRRVRRVRRVRRVRR 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53.5	49.1	171	12	071097
2	51.5	47.2	428	17	08TXS5
3	50	45.9	240	16	09PFU7
4	50	45.9	742	12	091PS3
5	49.5	44.0	349	10	09SHX2
6	48	44.0	241	17	08TW69
7	48	44.0	349	16	0987V7
8	48	44.0	842	15	09QMB3
9	47.5	43.6	421	10	098Y90
10	47	43.1	333	17	09Y8W8
11	47	43.1	331	17	08TXAS
12	47	43.1	760	12	099ARS
13	47	43.1	841	15	041556
14	47	43.1	862	15	090CG7
15	46.5	42.7	468	10	09FZM9
16	46	42.2	104	16	0986F0

17	46	42.2	201	2	09K523	Q9K523 mycobacteri
18	46	42.2	229	16	09WZL2	Q9WZL2 thermotoga
19	46	42.2	234	16	08U7B9	Q8U7B9 agrobacteri
20	46	42.2	242	9	094ML3	Q94ML3 bacterioph
21	46	42.2	263	10	09LNU0	Q9LNU0 arabidopsi
22	46	42.2	451	16	08XFW5	Q8XFW5 salmonella
23	46	42.2	452	16	08ZBM1	Q8ZBM1 yersinia pe
24	46	42.2	452	16	08X9F5	Q8X9F5 escherichia
25	46	42.2	673	16	08YDS8	Q8YDS8 bruceella me
26	46	42.2	723	12	09DUC4	Q9DUC4 tt virus. o
27	46	42.2	838	15	09DVL4	Q9DVL4 human immun
28	46	42.2	838	15	08UTC7	Q8UTC7 human immun
29	46	42.2	859	15	072940	Q72940 human immun
30	46	42.2	859	15	08UTD6	Q8UTD6 human immun
31	46	42.2	999	4	09NQ36	Q9NQ36 homo sapien
32	45.5	41.7	34	13	P83264	P83264 scomber sco
33	45	41.3	277	5	P92151	P92151 caenorhabdi
34	45	41.3	304	17	027855	027855 methanobact
35	45	41.3	316	2	033872	033872 xanthomonas
36	45	41.3	340	2	09L6V4	Q9L6V4 xanthomonas
37	45	41.3	394	17	08TVF5	Q8TVF5 methanopyru
38	45	41.3	405	1	08X259	Q8X259 methanopyru
39	45	41.3	577	2	049739	Q49739 mycobacteri
40	45	41.3	579	5	09W472	Q9W472 dirosophila
41	45	41.3	602	16	09CCU8	Q9CCU8 mycobacteri
42	45	41.3	650	5	095U14	Q95U14 dirosophila
43	45	41.3	847	15	0994M9	Q994M9 human immun
44	45	41.3	847	15	090CM2	Q90CM2 human immun
45	45	41.3	849	15	08UT64	Q8UT64 human immun

ALIGNMENTS

RESULT 1

071097 ID 071097 PRELIMINARY, PRT, 171 AA.

AC 071097/1

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE EVII.

OS Bovine adenovirus type 3 (Mastadenovirus boag).

OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.

OX NCBI_TaxID=10510;

ON [1]

RP SEQUENCE FROM N.A.

RC STRAIN=WB-1;

RX MEDLINE=98105785; PubMed=9445040;

RA Reddy P.S., Idamakanti N., Zakharichouk A.N., Baxi M.K., Lee J.B.,

RA Pyne C., Babluk L.A., Tikoo S.K.;

RT "Nucleotide sequence, genome organization, and transcription map of

RT bovine adenovirus type 3.,"

RL J. Virol. 72:1394-1402 (1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=WB-1;

RX MEDLINE=98118755; PubMed=9654686;

RA Baxi M.K., Reddy P.S., Zakharichouk A.N., Idamakanti N., Pyne C.,

RA Babluk L.A., Tikoo S.K.;

RT "Characterization of bovine adenovirus type 3 early region 2B.,"

RL Virus Genes 16:313-316 (1998).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=WB-1;

RX MEDLINE=98451815; PubMed=9778793;

RA Lee J.B., Baxi M.K., Idamakanti N., Reddy P.S., Zakharichouk A.N.,

RA Pyne C., Babluk L.A., Tikoo S.K.;

RT "Genetic organization and DNA sequence of early region 4 of bovine

RT adenovirus type 3.,"

RL Virus Genes 17:99-100 (1998).

RN [4]

RP SEQUENCE FROM N.A.

Mon Jun 9 12:26:57 2003

us-10-079-075-5.rsp

Page 8

Db 25 VRRIVRRIGTLARRVQ 42

Search completed: June 9, 2003, 11:56:37
Job time : 5.85106 secs

DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Proteamine YII (Clupeine YII).
 OS Clupea pallasi (Pacific herring), and
 OC Clupea harengus (Atlantic herring).
 OS Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
 CC Clupea.
 NCBI_TaxID=30724, 7950;
 OX NCBI [1]
 RN RP
 RP SPECIES=C.pallasi;
 RX MEDLINE=73223106; PubMed=4664740;
 RA Suzuki K., Ando T.;
 RT "Studies on proteamine, XVI. The complete amino acid sequence of
 RT clupeine YII.";
 RL J. Biochem. 72:1419-1432(1972).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=C.harengus;
 RA Chang W.J., Nukushina M., Ishii S., Nakahara C., Ando T.;
 RL Submitted (Aug-1970) to the PIR data bank.
 CC -1- FUNCTION: PROTEAMINE SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Nucleus.
 CC -1- TISSUE SPECIFICITY: TESTIS.
 DR PIR; A02677; CLHR2.
 DR PIR; A37575; CLHR2A.
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
 KW Testis; DNA condensation; Nucleic acid protein.
 SQ SEQUENCE 30 AA; 4049 MW; 7F9BBB0F3AD566 CRC64;
 QY
 Query Match 40.4%; Score 44; DB 1; Length 30;
 Best Local Similarity 58.3%; Pred. No. 2.1;
 Matches 14; Conservative 0; Mismatches 8; Indels 2; Gaps 1;
 QY 1 RRVRRVRRVRRVRRVRRVRR 24
 Db 2 RRRTRRASRVRR--RRPRVRSR 23
 RESULT 14
 VATE TREPA STANDARD; PRT; 232 AA.
 AC 083439;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE V-type ATP synthase subunit E (EC 3.6.3.14) (V-type ATPase subunit E).
 GN ATP6 OR TP0424.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 NCBI_TaxID=160;
 OX NCBI [1]
 RN RP
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Frazer C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
 RA McDonald L., Attlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 RT spirochete.";
 RL Science 281:375-388(1998).
 CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
 CC GRADIENT ACROSS THE MEMBRANE.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).

CC -1- SIMILARITY: BELONGS TO THE V-ATPASE E SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V01555; CAA24821.1; -.
 DR PIR; A03776; Q0834.
 DR PIR; S33026; S33026.
 DR InterPro; IPR003840; Herpes helicase.
 DR Pfam; PF02689; Herpes Helicase; 1.
 KW DNA replication; ATP-binding; Helicase; Early protein.
 FT NP_BIND 72 79
 SQ SEQUENCE 809 AA; 89853 MW; 434AA6EDAC01CC50 CRC64;
 QY
 Query Match 40.4%; Score 44; DB 1; Length 809;
 Best Local Similarity 50.0%; Pred. No. 59;
 Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 7 VRRVRRVRRVRRVRRVRR 24
 Db 11 RIVRARERARIVRARERAR 33
 RESULT 15
 HELI_EBV STANDARD; PRT; 809 AA.
 ID HELI_EBV
 AC P03214;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE Probable helicase.
 GN BBLF4.
 OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Lymphocryptovirus.
 OX NCBI_TaxID=10377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84270667; PubMed=6087149;
 RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
 RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Segun C.,
 RA Tufnell P.S., Watfull B.G.;
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
 RL Nature 310:207-211(1984).
 CC -1- FUNCTION: THIS PROTEIN MAY BE AN HELICASE AND IS REQUIRED FOR
 CC REPLICATION OF VIRAL DNA.
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL5,
 CC EBV-1 57, EBV BBLF4, HCMV UL105, AND VZV 55.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V01555; CAA24821.1; -.
 DR PIR; A03776; Q0834.
 DR PIR; S33026; S33026.
 DR InterPro; IPR003840; Herpes helicase.
 DR Pfam; PF02689; Herpes Helicase; 1.
 KW DNA replication; ATP-binding; Helicase; Early protein.
 FT NP_BIND 72 79
 SQ SEQUENCE 809 AA; 89853 MW; 434AA6EDAC01CC50 CRC64;

FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 644 644 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 823 823 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 863 AA; 97743 MW; B729CB5A6FAD1641 CRC64;

Query Match 41.3%; Score 45; DB 1; Length 863;
 Best Local Similarity 40.0%; Pred. No. 47;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 RVRVRVRVRVRVRVRVRVR 21
 DB 835 RILEIRRRARRVLRPRR 854

RESULT 11
 PRT2 TRUTH STANDARD; PRT; 34 AA.
 AC P02332;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-MAY-1992 (Rel. 22, Last annotation update)
 DE Protamines Z1 and Z2 (Thymnin Z1 and Z2).
 OS Thymnus thymus (Bluetin tuna).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;
 OC Scombridae; Thunnus.
 OC NCBI_TaxID=8237;
 RX MEDLINE=75039952; PubMed=4803475;
 RA Bretzel G.;
 RT "Thymnin, the protamine of the tuna fish: amino acid sequence of
 thymnin Z1. XIII. Communication on the structure of protamines from
 the studies of E. Waldschmidt-Leitz and coworkers.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:312-320 (1973).
 [2]
 RX MEDLINE=75039979; PubMed=4609881;
 RA Bretzel G.;
 RT "Thymnin, the protamine of the tuna fish: the amino acid sequence of
 thymnin Z2. XIV. Communication on the structure of the protamines
 described by E. Waldschmidt-Leitz et al.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:543-549 (1973).
 CC -1- FUNCTION: PROTIMINIS SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
 SPERM DURING THE HARPOID PHASE OF SPERMATOGENESIS. THEY COMPACT
 SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- MISCELLANEOUS: THE THYMNIN Z1 SEQUENCE IS SHOWN.
 DR PIR: A02664; TYTUZ1.
 DR PIR: A91657; TYTUZ2.
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
 KM Testis; DNA condensation; Nuclear protein.
 FT VARIANT 22 V->A (IN THYMNIN Z2).

SQ SEQUENCE 34 AA; 4641 MW; EEF72CED0158C3DE CRC64;

Query Match 40.8%; Score 44.5; DB 1; Length 34;
 Best Local Similarity 62.5%; Pred. No. 2.1;
 Matches 15; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 2 RVRVRVRVRVRVRVRVRVR 24
 DB 9 RVRVRVRVRVRVRVRVRVR 32

RESULT 12
 Y199 ARCFU STANDARD; PRT; 399 AA.
 AC 028380;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1899.
 GN AF1899.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OC NCBI_TaxID=2234;
 RX MEDLINE=98049343; PubMed=9389475;
 RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.U., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterback T.,
 RA Cotton P.W., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370 (1997).
 CC -1- SIMILARITY: BELONGS TO THE UPF0095 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: A800971; AAB89350.1; -.

DR TIGR: AF1899; -.

DR InterPro: IPR002936; DMAPrim_toprim.

DR Pfam: PF01751; Toprim; 1.

DR SMART: SM00493; TOPRIM; 1.

KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 399 AA; 44238 MW; F33233E0BBF57D8 CRC64;

Query Match 40.8%; Score 44.5; DB 1; Length 399;
 Best Local Similarity 26.5%; Pred. No. 25;
 Matches 9; Conservative 8; Mismatches 6; Indels 11; Gaps 1;

QY 1 RVRVRVRVRVRVRVRVRVR 23
 DB 127 KRIVERAMNLRHPEEPESERIVEIVQAIR 160
 RESULT 13
 PRT1 CLUPA STANDARD; PRT; 30 AA.
 AC P02335;

NADA METTH STANDARD; PRT; 304 AA.
 AC 027855;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Quinolinate synthetase A.
 GN NADA OR MTH1827.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteri; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=9803751; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Baehrizadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pochier B., Qiu D.,
 RA Spadafora R., Viscare R., Wang Y., Wierbowski J., Gibson R.,
 RA Jiwani N., Canuso A., Bush D., Safer H., Pietrowski S., Church G.M.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT delta: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- FUNCTION. Catalyzes the condensation of iminoaspartate with
 CC dihydroxyacetone phosphate to form quinolinate (By similarity).
 CC -1- PATHWAY: NAD biosynthesis; aspartate to NAMN; second step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE QUINOLINATE SYNTHETASE A FAMILY.
 CC SUBFAMILY 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; A800936; AAB6293.1; -
 DR InterPro; IPR003473; NADA.
 DR Pfam; PF02445; NADA; 1.
 DR TIGRFAMs; TIGR00550; nada; 1.
 KM Pyridine nucleotide biosynthesis; Complete proteome.
 SQ SEQUENCE 304 AA; 34393 MW; B5C48ACE4821430D CRC64;
 Query Match 41.3%; Score 45; DB 1; Length 304;
 Best Local Similarity 53.3%; Pred. No. 16;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
 RA Salanuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
 RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
 RT at risk for AIDS.";
 RL Science 232:1548-1553(1986).
 CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
 CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
 CC WAS PERINATALLY INFECTED BY HER MOTHER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M12507; AAB12990.1; -
 DR HIV; M12507; ENV5WMJ2.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane;
 KW Signal.
 FT SIGNAL. 1 29
 FT CHAIN 502 847
 FT DISULFID 53 73
 FT DISULFID 118 202
 FT DISULFID 125 193
 FT DISULFID 130 152
 FT DISULFID 215 244
 FT DISULFID 225 236
 FT DISULFID 293 326
 FT DISULFID 372 435
 FT DISULFID 379 408
 FT CARBOHYD 87 87
 FT CARBOHYD 134 134
 FT CARBOHYD 140 140
 FT CARBOHYD 151 151
 FT CARBOHYD 155 155
 FT CARBOHYD 183 183
 FT CARBOHYD 184 184
 FT CARBOHYD 194 194
 FT CARBOHYD 231 231
 FT CARBOHYD 238 238
 FT CARBOHYD 259 259
 FT CARBOHYD 273 273
 FT CARBOHYD 286 286
 FT CARBOHYD 292 292
 FT CARBOHYD 327 327
 FT CARBOHYD 334 334
 FT CARBOHYD 350 350
 FT CARBOHYD 356 356
 FT CARBOHYD 380 380
 FT CARBOHYD 386 386
 FT CARBOHYD 390 390
 FT CARBOHYD 400 400
 FT CARBOHYD 438 438
 FT CARBOHYD 450 450
 FT CARBOHYD 602 602
 FT CARBOHYD 607 607
 FT CARBOHYD 616 616
 FT CARBOHYD 628 628
 SQ SEQUENCE 847 AA; 96466 MW; CD1E33D73A45BCAE CRC64;
 Query Match 41.3%; Score 45; DB 1; Length 847;
 Best Local Similarity 35.0%; Pred. No. 46;
 Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;


```

OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaei A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki Y., Kuchida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
CC -1- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; APO00060; BAA80072.1; -
DR InterPro; IPR000054; Ribosomal_L31e.
DR Pfam; PF01198; Ribosomal_L31e.1.
DR PROSITE; PS01144; RIBOSOMAL_L31E; PALSE_NEG.
KW RIBOSOMAL protein; Complete proteome.
SQ SEQUENCE 105 AA; 12527 MW; 7E5DF7999E74A098 CRC64;

Query Match 47.2%; Score 51.5; DB 1; Length 105;
Best Local Similarity 57.7%; Pred. No. 0.75;
Matches 15; Conservative 2; Mismatches 6; Indels 3; Gaps 2;

Qy 1 RRVVRRVRRVRRVRRVRRVRR 24
Db 15 RRVVGRTRRRAL-RAVVRVREFVRR 39

RESULT 3
VCOT_ADE40
ID VCOT_ADE40 STANDARD; PRT; 185 AA.
AC Q89532;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
GN pVII.
OS Human adenovirus type 40.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28284;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RX MEDLINE=94087748; PubMed=8263936;
RA Davidson A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
RT "The DNA sequence of adenovirus type 40.";
RL J. Mol. Biol. 234:1308-1316(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dugan;
RX Pieniazek N.J., Slemenda S.B., Pieniazek D., Luftig R.B.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; L19443; AAC13963.1; -
DR EMBL; M86665; AAA42526.1; -
DR InterPro; IPR004912; Adeno_VII.
DR Pfam; PF03228; Adeno_VII.1.
FT PROPEP 1 23 BY SIMILARITY.
FT CHAIN 24 185 MAJOR CORE PROTEIN.
FT SITE 23 24 CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT (POTENTIAL).
SQ SEQUENCE 185 AA; 20518 MW; 4FB80E53EF218A9E CRC64;

Query Match 45.9%; Score 50; DB 1; Length 185;
Best Local Similarity 41.2%; Pred. No. 2.1;
Matches 14; Conservative 4; Mismatches 6; Indels 10; Gaps 1;

Qy 1 RRVVRRVRRVRRVRRVRRVRR 24
Db 99 RRLQRRRRPPTAMTAARAVLRAGRIGRRARR 132

RESULT 4
PRT ORYLA
ID PRT ORYLA STANDARD; PRT; 32 AA.
AC Q91185;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protamine.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorphia; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Orange-red; TISSUE=Testis;
RA Tamura M., Yamamoto H., Onitake K.;
RT "Cloning of protamine cDNA of the medaka (Oryzias latipes) and its
RT expression during spermatogenesis.";
RL Dev. Growth Differ. 36:419-425(1994).
CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Testis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D63796; BAA09865.1; -
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
SQ SEQUENCE 32 AA; 4220 MW; E10EC3A5B2ED6803 CRC64;

Query Match 42.2%; Score 46; DB 1; Length 32;
Best Local Similarity 70.0%; Pred. No. 1.2;
Matches 14; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy 5 RRVVRRVRRVRRVRRVRRVRR 24
Db 13 RRVVRRVRRVRRVRRVRRVRR 30

RESULT 5
PSS_ECOLI
ID PSS_ECOLI STANDARD; PRT; 451 AA.
AC P23830; P78256;
DT 01-NOV-1991 (Rel. 20, Created)

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:40:31 ; Search time 4.85106 Seconds
(without alignments)
205.199 Million cell updates/sec

Title: US-10-079-075-5

Perfect score: 109

Sequence: 1 RRVRRRVRRRVRRRVRRRVRR 24

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	47.7	33	1 PRTB_MUGCE	P08130 mugil cepha
2	51.5	47.2	105	1 RL31_AERPE	O99425 aeropyrum p
3	50	45.9	185	1 VCO7_ADE40	O89532 human adeno
4	46	42.2	32	1 PRT_ORYDA	O91185 oryzae lac
5	46	42.2	451	1 PSS_ECOLI	P23830 escherichia
6	45.5	41.7	70	1 L2M7_ADEB2	O96656 bovine aden
7	45	41.3	304	1 NADA_MERTH	O27885 methanobact
8	45	41.3	847	1 ENV_HV1W2	P05880 human immun
9	45	41.3	856	1 ENV_HV1W1	P31872 human immun
10	45	40.8	863	1 ENV_HV128	P05882 human immun
11	44.5	40.8	34	1 PRT2_THUTH	P02332 thunnus thy
12	44.5	40.8	399	1 Y199_ARCFU	O28380 archaeeoglob
13	44	40.4	30	1 PRT1_CLUPA	P02335 clupea pall
14	44	40.4	232	1 VATE_TREPA	O83439 treponema p
15	44	40.4	809	1 HELI_EBV	P03214 Epstein-Bar
16	44	40.4	898	1 SYA_MERTH	O27718 methanobact
17	43	39.4	30	1 PRTB_ONCMY	P12819 oncorhynch
18	43	39.4	170	1 YACP_BACSU	P37574 bacillus su
19	43	39.4	266	1 IF2A_SULTO	O97279 sulfobobus
20	42.5	39.0	133	1 VCO7_ADE04	O26820 methanobact
21	42.5	39.0	133	1 VCO7_ADE04	O96831 human adeno
22	42.5	39.0	526	1 EX7L_RHIME	O92877 rhizobium m
23	42	38.5	30	1 PRT3_ONCMY	P02332 oncorhynch
24	42	38.5	30	1 PRT4_ONCMY	P02333 oncorhynch
25	42	38.5	31	1 PRT2_CLUPA	P02336 clupea pall
26	42	38.5	36	1 Y4KD_BPCBP	P19188 bacterioph
27	42	38.5	75	1 RL29_PYRAE	O82711 pyrobaculum
28	42	38.5	99	1 VHSB_BPT7	P03701 bacterioph
29	42	38.5	263	1 IF2A_SULTO	O97360 sulfobobus
30	42	38.5	317	1 RPSD_STRAU	B27785 streptomyces
31	42	38.5	346	1 RUVB_BROME	O84155 bruceella me
32	42	38.5	396	1 HRDA_STRIC	P18182 streptomyces
33	42	38.5	761	1 MAO2_RHIME	O30808 rhizobium m

34	41.5	38.1	34	1 PRT_DICLA	O99627 dicentrarch
35	41.5	38.1	34	1 PRT_PERYV	P29629 perca flav
36	41.5	38.1	87	1 SSSI_SCYCA	P13275 scyllorhinu
37	41.5	38.1	323	1 CC39_CAEEL	O09455 caenorhabdi
38	41.5	38.1	358	1 VCO7_ADE40	P48753 human adeno
39	41	37.6	202	1 Y677_TREPA	O83683 treponema p
40	41	37.6	294	1 E434_ADE02	P03239 human adeno
41	41	37.6	308	1 YFER_ECOLI	P77500 escherichia
42	41	37.6	846	1 ENV_HV1ND	P18799 human immun
43	40.5	37.2	34	1 PRT1_SAROR	P25327 sarda orien
44	40.5	37.2	34	1 PRT1_THUTH	P02321 thunnus thy
45	40.5	37.2	58	1 HSP3_HORSE	P15343 equus caball

ALIGNMENTS

RESULT 1

ID	PRTB_MUGCE	STANDARD	PRT	33 AA.
AC	P08130			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Protamine M6/M7 (Mugiline beta).			
OS	Mugil cephalus (Flathead mullet).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Mugilomorpha; Mugilidae;			
OC	Mugil.			
NC	NCBI_TaxID=48193;			
RN	[1]			
RP	SEQUENCE:			
RX	MEDLINE=87279669; PubMed=3301825;			
RA	Okamoto Y., Muta E., Ota S.			
RT	Primary structures of M6 and M7 of mugiline beta (Mugil japonicus)."			
RL	J. Biochem. 101:1017-1024(1987).			
CC	-1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF			
CC	SPEM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPLE			
CC	SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- TISSUE SPECIFICITY: TESTIS.			
CC	-1- MISCELLANEOUS: THE SEQUENCE OF COMPONENT M6 IS SHOWN.			
DR	PIR: A26762. 426762.			
DR	PIR: B26762. B26762.			
KW	Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;			
KW	Testis; DNA condensation; Nuclear protein.			
FT	VARIANT 6 6 E -> Q (IN M7 COMPONENT).			
FT	VARIANT 22 22 I -> M.			
SQ	SEQUENCE 33 AA; 4473 MW; 48407DE638A0D29E CRC64;			

Query Match 47.7%; Score 52; DB 1; Length 33;
Best Local Similarity 56.5%; Pred. No. 0.2;
Matches 13; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

DB 9 RRVRRRVRRRVRRRVRRRVRR 24

RESULT 2

ID	RL31_AERPE	STANDARD	PRT	105 AA.
AC	O99425			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	50S ribosomal protein L31e.			
GN	RP131E OR APE1087.			
OS	Aeropyrum pernix.			
OC	Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;			
OC	Desulfurococcaceae; Aeropyrum.			

Science 294, 2323-2328, 2001
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*
A:Reference number: A97359; PMID:11743194
A:Accession: G98173
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK8913.1; PID:G15158686; GSPDB:GN00170
A:Gene: AGR_L_676
A:Map position: linear chromosome

Query Match 42.2%; Score 46; DB 2; Length 234;
Best Local Similarity 34.8%; Pred. No. 46;
Matches 8; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 2 RVRVRVRVRVRVRVRVRVRVR 24
DB 136 RVSEAIRNVRPAIEIRILKR 158

RESULT 13

E86215
protein T6D22.16 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E86215
R:Theologas, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.N.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, C.W.; Hughes, B.; Hultzer, L.
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matzl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakao, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86215
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <STO>
A:Cross-references: GB:AE005172; NID:98778838; PIDN:AAF79837.1; GSPDB:GN00141
A:Gene: T6D22.16
A:Map position: 1

Query Match 42.2%; Score 46; DB 2; Length 263;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 VRRVRVRVRVRVRVRVRVR 22
DB 218 VRRVRVRVRVRVRVRVRVR 237

RESULT 14

AE0831
CDPdiacylglycerol-serine O-phosphatidyltransferase (EC 2.7.8.8) - *Salmonella enterica* serovar Typhimurium
A:Species: *Salmonella enterica* subsp. *enterica* serovar Typhimurium
A>Note: This species has also been called *Salmonella typhi*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 03-Jun-2002
C:Accession: AE0831
R:Perkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, F.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A>Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0831
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-451 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05836.1; PID:G16503811; GSPDB:GN00176
A:Gene: STY2845
A:Superfamily: Escherichia coli CDPdiacylglycerol-serine O-phosphatidyltransferase
C:Keywords: transferase

Query Match 42.2%; Score 46; DB 2; Length 451;
Best Local Similarity 36.4%; Pred. No. 83;
Matches 8; Conservative 10; Mismatches 2; Indels 2; Gaps 1;

QY 1 RRVRRVRVRVRVRVRVRVR 22
DB 432 RKLIRLRIRI--RIDRLIRIL 451

RESULT 15

H65036
CDPdiacylglycerol-serine O-phosphatidyltransferase (EC 2.7.8.8) - *Escherichia coli* (serotype O157:H7)
N/Alternate names: phosphatidylserine synthase
C:Species: *Escherichia coli*
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: H65036; JH0368; A40406
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C. A.; Rose, D.J.; Mau, B.; Shao, Y.
A:Title: The complete genome sequence of *Escherichia coli* K-12.
Science 277, 1453-1462, 1997
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H65036
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-452 <BLAT>
A:Cross-references: GB:AE000345; GB:U00096; NID:91788939; PIDN:AACT5638.1; PID:G1788940, R:Experimental source: strain K-12, substrain MG1655
R:DeChavigny, A.; Heacock, P.N.; Dowhan, W.
J. Biol. Chem. 266, 5333-5332, 1991
A>Title: Sequence and inactivation of the *psa* gene of *Escherichia coli*: phosphatidylethanolamine transferase
A:Reference number: JH0368; MUID:91161632; PMID:2002065
A:Accession: JH0368
A:Molecule type: DNA
A:Residues: 2-32, 'R', '34-78', 'DD', '80-165', 'NIA', '169-287', 'FV', '290-309', 'S', '311-452' <DEC>
A:Cross-references: GB:M58699; NID:9147388; PIDN:AA97504.1; PID:G147389
R:DeChavigny, A.; Heacock, P.N.; Dowhan, W.
J. Biol. Chem. 266, 10710, 1991
A:Reference number: A40406; MUID:91244856; PMID:2037609
A:Contents: erratum
A:Accession: A40406
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 2-32, 'R', '34-78', 'DD', '80-165', 'NIA', '169-287', 'FV', '290-309', 'S', '311-452' <DEC>
C:Comment: The enzyme catalyzes the committed step to phosphatidylethanolamine biosynthesis

Query Match 42.2%; Score 46; DB 1; Length 452;
Best Local Similarity 36.4%; Pred. No. 84;
Matches 8; Conservative 10; Mismatches 2; Indels 2; Gaps 1;

QY 1 RRVRRVRVRVRVRVRVRVR 22
DB 433 RKLIRLRIRI--RIDRLIRIL 452

Search completed: June 9, 2003, 12:03:09
Job time: 10.9574 secs

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: D72484

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-323 <KAM>

A:Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BA081532.1; PID:dl045318; PID:9510

A:Experimental source: strain K1

C:Genetics:

A:Gene: AB2516

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology F:29-221/Domain: ATP-binding cassette homology <ABC>

Query Match 43.1%; Score 47; DB 2; Length 323;

Best Local Similarity 42.9%; Pred. No. 46;

Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 RVRVRVRVRVRVRVRVRV 22

DB 179 KARVRVSRIRVRVRVTL 199

RESULT 8

S27771

RNA-directed DNA polymerase (EC 2.7.7.49) - African malaria mosquito transposon RT1 (fr

N:Alternate names: reverse transcriptase

C:Species: Anopheles gambiae (African malaria mosquito)

C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Sep-1997

C:Accession: S27771

R:Residues: 1-1212 <BES>

A:Cross-references: EMBL:M33690; NID:9159615; PID:9159617

A:Experimental source: strain K1

C:Genetics:

A:Gene: S27771

A:Molecule type: DNA

A:Residues: 1-1212 <BES>

A:Cross-references: EMBL:M33690; NID:9159615; PID:9159617

C:Keywords: nucleotidyltransferase

Query Match 42.7%; Score 46.5; DB 2; Length 1212;

Best Local Similarity 54.5%; Pred. No. 1.8e+02;

Matches 12; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 1 RVRVRVRVRVRVRVRVRV 21

DB 871 RRLRVRVRVRVRVRVRV 892

RESULT 9

I51089

protamine - Japanese medaka

C:Species: Oryzias latipes (Japanese medaka)

C>Date: 11-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 19-May-2000

C:Accession: I51089

R:Tamura, M.; Yamamoto, H.; Onitake, K.

Dev. Growth Differ. 36, 419-425, 1994

A:Title: Cloning of protamine cDNA of the medaka (*Oryzias latipes*) and its expression dur

A:Reference number: I51089

A:Accession: I51089

A>Status: preliminary; translated from GB/EMBL/DDB

A:Molecule type: mRNA

A:Residues: 1-32 <TAM>

A:Cross-references: GB:D63796; NID:9961485; PIDN:BA09865.1; PID:9961486

C:Superfamily: protamine Y2.

Query Match 42.2%; Score 46; DB 2; Length 32;

Best Local Similarity 70.0%; Pred. No. 7.7;

Matches 14; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 5 RVRVRVRVRVRVRVRVRV 24

DB 13 RVRVRVRVRVRVRVRVGR 30

RESULT 10

G72337

hypothetical protein TW0753 - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: G72337

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: G72337

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229 <ARN>

A:Cross-references: GB:AE001745; GB:AE000512; NID:94981278; PIDN:AAD35835.1; PID:9498128

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TW0753

C:Superfamily: spore germination protein C2; bioc homology

F:48-147/Domain: bioc homology <BIOC>

Query Match 42.2%; Score 46; DB 2; Length 229;

Best Local Similarity 55.6%; Pred. No. 45;

Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RVRVRVRVRVRVRVRV 18

DB 127 RVRVRVRVRVRVRVRV 144

RESULT 11

AF3113

transcription regulator, TetR family Atu4530 [imported] - Agrobacterium tumefaciens (str

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AF3113

R:Wood, D.W.; Seubald, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J

erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kuyavain, T.; Levy, R.; Li, M.; McClell

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AF3113

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-230 <KUR>

A:Cross-references: GB:AE008689; PIDN:AAL45324.1; PID:917743015; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu4530

A:Map position: linear chromosome

Query Match 42.2%; Score 46; DB 2; Length 230;

Best Local Similarity 34.8%; Pred. No. 45;

Matches 8; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 2 RVRVRVRVRVRVRVRVRV 24

DB 132 RVSEAIRVNVNPAIEIRILKR 154

RESULT 12

G96173

probable transcription regulator PA0243 [imported] - Agrobacterium tumefaciens (strain

C:Species: Agrobacterium tumefaciens

C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C:Accession: G96173

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Orozco, B.; Goldman

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.

probable ribosomal protein L31 APE1087 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 02-Aug-2002
C:Accession: H72708
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; PMID:9310339; PMID:10382966
A:Accession: H72708
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-105 <RAW>
A:Cross-references: DDBJ:AE00060; NID:95104188; PIDN:BAA80072.1; PID:d1043858; PID:9510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1087
C:Superfamily: rat ribosomal protein L31

Query Match 47.2%; Score 51.5; DB 2; Length 105;
Best Local Similarity 57.7%; Pred. No. 4.6;
Matches 15; Conservative 2; Mismatches 6; Indels 3; Gaps 2;

QY 1 RRVV--RRRVRRRVRRRVRRRV 24
DB 15 RRVVGRTRRRAL-RAVVRVREFFVR 39

RESULT 4

F82790 GMP synthase XF0560 [imported] - Xylella fastidiosa (strain 9asc)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82790
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; PMID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82790
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-240 <STM>
A:Cross-references: GB:AE003903; GB:AE003849; NID:9105416; PIDN:AAF83370.1; GSPDB:GN001
A:Experimental source: strain 9asc
R:Stimpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.; B
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H
as-Neto, E.; Docena, C.; El-Dorcy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitzjima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigz
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshakko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0560

Query Match 45.9%; Score 50; DB 2; Length 240;
Best Local Similarity 46.4%; Pred. No. 15;
Matches 13; Conservative 4; Mismatches 5; Indels 6; Gaps 1;

QY 1 RRVVRRV-----RRVRRVRRVRRV 22
DB 213 RRVVRRVRRVRRVRRVRRVRRVRRV 240

RESULT 5

C84477

hypothetical protein AC2906420 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84477
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; PMID:20083487; PMID:10617197
A:Accession: C84477
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <STO>
A:Cross-references: GB:AE002093; NID:94646226; PIDN:AAD26890.1; GSPDB:GN00139
C:Genetics:
A:Gene: AC2906420
A:Map position: 2

Query Match 45.4%; Score 49.5; DB 2; Length 349;
Best Local Similarity 48.0%; Pred. No. 24;
Matches 12; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 RRVVRRVRRVRRVRRVRRVRRVRRV 24
DB 278 RRVVRRVRRVRRVRRVRRVRRVRRV 302

RESULT 6

D96641 hypothetical protein T25B24.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96641
R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chn, C.W.; Hughes, B.; Hultzar, L.
ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maitl, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; PMID:21016719; PMID:11130712
A:Accession: D96641
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-421 <STO>
A:Cross-references: GB:AE005173; NID:94585877; PIDN:AAD25550.1; GSPDB:GN00141
C:Genetics:
A:Gene: T25B24.5
A:Map position: 1

Query Match 43.6%; Score 47.5; DB 2; Length 421;
Best Local Similarity 45.7%; Pred. No. 51;
Matches 16; Conservative 3; Mismatches 3; Indels 13; Gaps 2;

QY 3 VRRVRRVRRVRRVRRVRRVRRVRRV 24
DB 336 VRRVRRVRRVRRVRRVRRVRRVRRV 370

RESULT 7

D72484 Probable ATP-binding protein APE2516 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 17-Mar-2000
C:Accession: D72484
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:48:06 / Search time 9.95745 Seconds

(without alignments)
231.709 Million cell updates/sec

Title: US-10-079-075-5

Sequence: 1 RRVRRVRRVRRVRRVRRVRR 24

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR 73: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	47.7	33	2 A26762	protamine (mugilin
2	52	47.7	33	2 B26762	probable ribosomal
3	51.5	47.2	105	2 H72708	GMP synthase XR056
4	50	45.9	240	2 F82790	hypothetical prote
5	49.5	45.4	349	2 C84477	hypothetical prote
6	47.5	43.6	421	2 D96641	probable ATP-bindi
7	47	43.1	333	2 D72484	RNA-directed DNA p
8	46.5	42.7	1212	2 S27771	protamine - Japane
9	46	42.2	32	2 I51089	hypothetical prote
10	46	42.2	229	2 G72337	transcription regu
11	46	42.2	230	2 AF3113	probable transcrip
12	46	42.2	234	2 G98173	protein T6D22.16 l
13	46	42.2	263	2 E86215	CDPdiacylglycerol-
14	46	42.2	451	2 AE0831	CDPdiacylglycerol-
15	46	42.2	452	1 H65036	CDPdiacylglycerol-
16	46	42.2	452	2 AG0397	CDPdiacylglycerol-
17	46	42.2	452	2 B85905	CDPdiacylglycerol-
18	46	42.2	452	2 D91060	CDPdiacylglycerol-
19	46	42.2	673	2 AG3521	CDPdiacylglycerol-
20	45.5	41.7	34	2 JN0582	protamine (secombr
21	45	41.3	277	2 T27597	hypothetical prote
22	45	41.3	304	2 C69111	guinolate synthase
23	45	41.3	377	2 S72834	hypothetical prote
24	45	41.3	602	2 E86958	probable transport
25	45	41.3	856	1 VCLP3W	env polyprotein pr
26	45	41.3	2351	2 G71415	hypothetical prote
27	44.5	40.8	34	1 TYTU21	protamine 21 - blu
28	44.5	40.8	399	2 B69487	conserved hypotet
29	44.5	40.8	601	2 T36323	probable membrane

30	44	40.4	30	1 CLHRV2	protamine YII - Pa
31	44	40.4	30	1 CLHR2A	protamine YII - At
32	44	40.4	58	2 H91110	hypothetical prote
33	44	40.4	232	2 E71325	probable V-type AT
34	44	40.4	276	2 C75508	hypothetical prote
35	44	40.4	294	2 D83108	hypothetical prote
36	44	40.4	374	2 H91251	probable tail prot
37	44	40.4	382	2 T34940	ATP dependent heli
38	44	40.4	779	2 AF1094	ATP dependent heli
39	44	40.4	809	1 QOBR34	ATP dependent heli
40	44	40.4	898	2 A69092	alanine-tRNA ligase
41	44	40.4	1153	2 A97179	ATP-dependent exon
42	44	40.4	1677	2 T14267	Xin protein, stage
43	44	40.4	2049	2 T29227	hypothetical prote
44	43.5	39.9	317	2 B75012	methanol dehydrog
45	43.5	39.9	1082	2 H81982	hypothetical prote

ALIGNMENTS

RESULT 1

A26762
protamine (mugiline beta) M6 - Formosan gray mullet
C/Species: Mugil japonicus (Formosan gray mullet)
C/Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #ext_change 23-Feb-1997
C/Accession: A26762
R/Okamoto, Y.; Muta, E.; Ota, S.
J. Biochem. 101, 1017-1024, 1987
A/Title: Primary structures of M6 and M7 of mugiline beta-(Mugil japonicus).
A/Reference number: A91909; PMID:8727969; PMID:3301825
A/Accession: A26762
A/Molecule type: protein
A/Residues: 1-33 <OKA>
C/Superfamily: protamine Y2
C/Keywords: DNA binding, nucleus

Query Match 47.7%; Score 52; DB 2; Length 33;
Best Local Similarity 56.5%; Pred. No. 1.4;
Matches 13; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 RVRVRVRRVRRVRRVRRVRR 24
DB 9 RPIRRRRRRAPIRRRRRVRR 31

RESULT 2

B26762
protamine (mugiline beta) M7 - Formosan gray mullet
C/Species: Mugil japonicus (Formosan gray mullet)
C/Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #ext_change 23-Feb-1997
C/Accession: B26762
R/Okamoto, Y.; Muta, E.; Ota, S.
J. Biochem. 101, 1017-1024, 1987
A/Title: Primary structures of M6 and M7 of mugiline beta-(Mugil japonicus).
A/Reference number: A91909; PMID:8727969; PMID:3301825
A/Accession: B26762
A/Molecule type: protein
A/Residues: 1-33 <OKA>
C/Superfamily: protamine Y2
C/Keywords: DNA binding, nucleus

Query Match 47.7%; Score 52; DB 2; Length 33;
Best Local Similarity 56.5%; Pred. No. 1.4;
Matches 13; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 RVRVRVRRVRRVRRVRRVRR 24
DB 9 RPIRRRRRRAPIRRRRRVRR 31

RESULT 3
H72708

ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796,899
FILING DATE: 06-FEB-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/319,544
FILING DATE: 07-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 028754-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 515 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-796-899-23

Query Match 59.3%; Score 32; DB 4; Length 515;
Best Local Similarity 63.6%; Pred. No. 2.9e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVVRVVRVVR 11
DB 314 KVVAVQVRVR 324

RESULT 15
US-08-022-835-6
Sequence 6, Application US/08022835
Patent No. 5420030
GENERAL INFORMATION:
APPLICANT: Reitz Jr., Marvin S.
APPLICANT: Franchini, Genoveffa
APPLICANT: Markham, Phillip D.
APPLICANT: Gallo, Robert C.
APPLICANT: Lori, Franco C.
APPLICANT: Popovic, Mikulas
APPLICANT: Gartner, Suzanne
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: Eleventh Floor, 1615 L. Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/022,835
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-022-835-6

Query Match 59.3%; Score 32; DB 1; Length 855;
Best Local Similarity 54.5%; Pred. No. 4.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVVRVVRVVR 11
DB 827 RVIEVLQRAVR 837

Search completed: June 9, 2003, 12:04:51
Job time: 6.08511 secs

NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,682
FILING DATE: 18-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/786,748
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-705-5000
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5945507e
US-08-932-682-55

Query Match 59.3%; Score 32; DB 2; Length 21;
Best Local Similarity 45.5%; Pred. No. 13;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVVRVVRVVR 11
|:|:|:|:|:|
DB 1 RRRPRRRR 11

RESULT 12
US-08-846-762-46
Sequence 46, Application US/08846762A
Patent No. 5994072
GENERAL INFORMATION:
APPLICANT: Lam, Joseph S.
APPLICANT: Burrows, Lori
APPLICANT: Charter, Deborah
APPLICANT: de Kievit, Teresa
TITLE OF INVENTION: No. 5994072e1 Proteins Involved in the Synthesis and Assembly
FILE REFERENCE: 6580-089
CURRENT APPLICATION NUMBER: US/08/846,762A
CURRENT FILING DATE: 1997-04-30
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 46
LENGTH: 35
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-08-846-762-46

Query Match 59.3%; Score 32; DB 2; Length 35;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVVRVVR 10
|:|:|:|:|:|

DB 26 VIAVRRVVR 34

RESULT 13
US-09-088-425-2
Sequence 2, Application US/09088425
Patent No. 6171843
GENERAL INFORMATION:
APPLICANT: BANDMAN, OLGA
APPLICANT: LAL, PREETI
APPLICANT: CORLEY, NEIL C.
APPLICANT: PATTERSON, CHANDRA
APPLICANT: BAUGHN, MARIN R.
TITLE OF INVENTION: HUMAN ISOMERASE-LIKE PROTEINS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,425
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0529 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 443 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TWLR3D701
CLONE: 289973
US-09-088-425-2

Query Match 59.3%; Score 32; DB 4; Length 443;
Best Local Similarity 87.5%; Pred. No. 2,5e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RVVRVVRVVR 11
|:|:|:|:|:|
DB 7 RVVRVVRVVR 14

RESULT 14
US-08-796-899-23
Sequence 23, Application US/08796899
Patent No. 6160202
GENERAL INFORMATION:
APPLICANT: BUSTOS, Mauricio M
APPLICANT: CHERN, Maw-Sheng
TITLE OF INVENTION: MODIFICATION OF SEED CROPS WITH
TRANSCRIPTION FACTORS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-55

Query Match 59.3%; Score 32; DB 1; Length 21;
Best Local Similarity 45.6%; Pred. No. 13;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVVRVVRVVR 11
|:|:|:|:|:
DB 1 RVRPRRIR 11

RESULT 9
US-08-932-682-38
Sequence 38, Application US/08932682
Patent No. 5945507
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mietzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,682
FILING DATE: 18-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/786,748
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-705-5000
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acid
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5945507e
US-08-932-682-38

Query Match 59.3%; Score 32; DB 2; Length 21;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVVRVVRVVR 11
|:|:|:|:|:
DB 1 RVRPRRIR 11

RESULT 10
US-08-932-682-39
Sequence 39, Application US/08932682
Patent No. 5945507
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mietzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,682
FILING DATE: 18-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/786,748
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-705-5000
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5945507e
US-08-932-682-39

Query Match 59.3%; Score 32; DB 2; Length 21;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVVRVVRVVR 11
|:|:|:|:|:
DB 1 RVRPRRIR 11

RESULT 11
US-08-932-682-55
Sequence 55, Application US/08932682
Patent No. 5945507
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mietzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-849-486-10

Query Match 59.3%; Score 32; DB 3; Length 15;
Best Local Similarity 50.0%; Pred. No. 9.3;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVRVVRVVR 12
|:|:|:|:|:
DB 1 RLRRLRLRLR 12

RESULT 6
US-08-786-748A-38
Sequence 38, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mletzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-38

Query Match 59.3%; Score 32; DB 1; Length 21;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVRVVRVVR 11
|:|:|:|:|:
DB 1 RVRVVRGACR 11

RESULT 7
US-08-786-748A-39
Sequence 39, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.

APPLICANT: Tencza, Sarah B.
APPLICANT: Mletzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-39

Query Match 59.3%; Score 32; DB 1; Length 21;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVRVVRVVR 11
|:|:|:|:|:
DB 1 RVRVVRGACR 11

RESULT 8
US-08-786-748A-55
Sequence 55, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mletzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,748A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/010,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2500
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5714577e
US-08-786-748A-30

Query Match 61.1%; Score 33; DB 1; Length 17;
Best Local Similarity 50.0%; Pred. No. 7.2;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RVRVVRVRRVRR 12
| : : : : :
Db 4 RACRAIRRRPR 15

RESULT 3
US-08-932-682-30
Sequence 30, Application US/08932682
Patent No. 5945507
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Mietzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTS, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,682
FILING DATE: 18-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/786,748
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rochelle K. Seide
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: AP30421-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-705-5000
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5945507e
US-08-932-682-30

Query Match 61.1%; Score 33; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 7.2;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RVRVVRVRRVRR 12
| : : : : :
Db 4 RACRAIRRRPR 15

RESULT 4
US-08-846-762-93
Sequence 93, Application US/08846762A
Patent No. 5994072
GENERAL INFORMATION:
APPLICANT: Lam, Joseph S.
APPLICANT: Burrows, Lori
APPLICANT: Charter, Deborah
APPLICANT: de Kievit, Teresa
TITLE OF INVENTION: OF O-Antigen in Pseudomonas Aeruginosa
FILE REFERENCE: 6580-089
CURRENT APPLICATION NUMBER: US/08/846,762A
CURRENT FILING DATE: 1997-04-30
NUMBER OF SEQ ID NOS: 100
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 93
LENGTH: 355
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-08-846-762-93

Query Match 61.1%; Score 33; DB 2; Length 355;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVRVVRVRRVRR 12
| : : : : :
Db 337 RITRWVRRMKRR 348

RESULT 5
US-08-849-486-10
Sequence 10, Application US/08849486
Patent No. 6080724
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
NUMBER OF SEQUENCES: 10
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,486
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95 11714
FILING DATE: 05-OCT-1995
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:50:31 ; Search time 4.08511 Seconds
(without alignments)
86.430 Million cell updates/sec

Title: US-10-079-075-4

Sequence: 1 RVVRRVRRVRR 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	68.5	320	2	US-08-846-762-16
2	33	61.1	17	1	US-08-786-748A-30
3	33	61.1	17	2	US-08-932-682-30
4	33	61.1	355	3	US-08-846-762-93
5	32	59.3	15	3	US-08-849-486-10
6	32	59.3	21	1	US-08-786-748A-38
7	32	59.3	21	1	US-08-786-748A-39
8	32	59.3	21	1	US-08-786-748A-55
9	32	59.3	21	2	US-08-932-682-38
10	32	59.3	21	2	US-08-932-682-39
11	32	59.3	21	2	US-08-932-682-55
12	32	59.3	35	2	US-08-846-762-46
13	32	59.3	443	4	US-09-088-425-2
14	32	59.3	515	4	US-08-786-899-23
15	32	59.3	855	1	US-08-022-835-6
16	32	59.3	855	1	US-08-388-809-6
17	32	59.3	855	2	US-08-647-714-6
18	32	59.3	1124	4	US-09-191-786-1
19	32	59.3	5087	4	US-09-144-085-1
20	32	59.3	6095	4	US-09-144-085-2
21	31	57.4	15	1	US-08-179-632-22
22	31	57.4	15	1	US-08-440-174A-22
23	31	57.4	15	5	PCT-US95-00062-22
24	31	57.4	31	1	US-08-179-632-9
25	31	57.4	31	1	US-08-440-174A-9
26	31	57.4	31	5	PCT-US95-00062-9
27	31	57.4	40	1	US-08-179-632-3

28	31	57.4	40	1	US-08-440-174A-3	Sequence 3, Appl1
29	31	57.4	40	5	PCT-US95-00062-3	Sequence 3, Appl1
30	31	57.4	103	3	US-09-191-647-6	Sequence 6, Appl1
31	31	57.4	103	3	US-09-540-245A-6	Sequence 6, Appl1
32	31	57.4	103	4	US-09-540-153-6	Sequence 6, Appl1
33	31	57.4	262	1	US-08-403-379A-1	Sequence 1, Appl1
34	31	57.4	262	2	US-08-929-414-1	Sequence 1, Appl1
35	31	57.4	263	2	US-08-557-309B-51	Sequence 51, Appl1
36	31	57.4	263	3	US-08-834-306-51	Sequence 51, Appl1
37	31	57.4	263	4	US-08-993-674A-51	Sequence 51, Appl1
38	31	57.4	263	4	US-09-256-976-51	Sequence 51, Appl1
39	31	57.4	264	2	US-07-857-224B-24	Sequence 24, Appl1
40	31	57.4	334	1	US-08-287-442-9	Sequence 9, Appl1
41	31	57.4	334	1	US-08-459-701-9	Sequence 9, Appl1
42	31	57.4	334	1	US-08-460-298-9	Sequence 9, Appl1
43	31	57.4	334	1	US-08-459-174-9	Sequence 9, Appl1
44	31	57.4	424	2	US-08-715-568A-1	Sequence 1, Appl1
45	31	57.4	442	3	US-08-834-306-52	Sequence 52, Appl1

ALIGNMENTS

RESULT 1
US-08-846-762-16
Sequence 16, Application US/08846762A
Patent No. 5994072
GENERAL INFORMATION:
APPLICANT: Lam, Joseph S.
APPLICANT: Burrows, Lori
APPLICANT: Charter, Deborah
TITLE OF INVENTION: No. 5994072el Proteins Involved in the Synthesis and Assembly
TITLE OF INVENTION: of O-Antigen in Pseudomonas Aeruginosa
FILE REFERENCE: 6580-089
CURRENT APPLICATION NUMBER: US/08/846, 762A
CURRENT FILING DATE: 1997-04-30
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 320
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-08-846-762-16

Query Match 68.5%; Score 37; DB 2; Length 320;
Best Local Similarity 72.7%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRRVRRVRR 12
|: |||||
Db 28 VVRRVRRVRR 38

RESULT 2
US-08-786-748A-30
Sequence 30, Application US/08786748A
Patent No. 5714577
GENERAL INFORMATION:
APPLICANT: Ronald, Montelaro C.
APPLICANT: Tencza, Sarah B.
APPLICANT: Metzner, Timothy A.
TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSES: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

```
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-7

Query Match          100.0%; Score 54; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RVRVVRVRRVRR 12
Db      13 RVRVVRVRRVRR 24

RESULT 13
US-09-785-059-8
; Sequence 8, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-8

Query Match          100.0%; Score 54; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RVRVVRVRRVRR 12
Db      1 RVRVVRVRRVRR 12

RESULT 14
US-10-079-075-8
; Sequence 8, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-8

Query Match          100.0%; Score 54; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RVRVVRVRRVRR 12
Db      1 RVRVVRVRRVRR 12

RESULT 15
US-09-785-058-8
; Sequence 8, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Metzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-8

Query Match          100.0%; Score 54; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RVRVVRVRRVRR 12
Db      1 RVRVVRVRRVRR 12

Search completed: June 9, 2003, 12:34:08
Job time : 7.38298 secs
```

FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-6

Query Match 100.0%; Score 54; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRR 12
DB 7 RRVRRVRRVRR 18

RESULT 8
US-10-079-075-6
Sequence 6, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-6

Query Match 100.0%; Score 54; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRR 12
DB 7 RRVRRVRRVRR 18

RESULT 9
US-09-785-058-6
Sequence 6, Application US/09785058
Publication No. US20030036627A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A 34001 / 072396.0222
CURRENT APPLICATION NUMBER: US/09/785.058
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-6

Query Match 100.0%; Score 54; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRR 12
DB 7 RRVRRVRRVRR 18

RESULT 10
US-09-785-059-7
Sequence 7, Application US/09785059
Patent No. US20020163279A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A33577 / 072396.0217
CURRENT APPLICATION NUMBER: US/09/785.059
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 42
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-7

Query Match 100.0%; Score 54; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRR 12
DB 13 RRVRRVRRVRR 24

RESULT 11
US-10-079-075-7
Sequence 7, Application US/10079075
Publication No. US20020188102A1
GENERAL INFORMATION:
APPLICANT: Ronald C. Montelaro
APPLICANT: Timothy A. Metzner
TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
FILE REFERENCE: A34001-A / 072396.0222
CURRENT APPLICATION NUMBER: US/10/079.075
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 42
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-7

Query Match 100.0%; Score 54; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVRRVRRVRR 12
DB 13 RRVRRVRRVRR 24

RESULT 12
US-09-785-058-7
Sequence 7, Application US/09785058
Publication No. US20030036627A1

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-4
```

```
Query Match          100.0%; Score 54; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00092;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 RVRVVRVRVRR 12
        |||||
Db      1 RVRVVRVRVRR 12
```

```
RESULT 3
US-09-785-058-4
; Sequence 4, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-4
```

```
Query Match          100.0%; Score 54; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00092;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 RVRVVRVRVRR 12
        |||||
Db      1 RVRVVRVRVRR 12
```

```
RESULT 4
US-09-785-059-5
; Sequence 5, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09/785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-5
```

```
Query Match          100.0%; Score 54; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 RVRVVRVRVRR 12
        |||||
Db      13 RVRVVRVRVRR 24
```

```
RESULT 5
US-10-079-075-5
; Sequence 5, Application US/10079075
; Publication No. US20020188102A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-10-079-075-5
```

```
Query Match          100.0%; Score 54; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 RVRVVRVRVRR 12
        |||||
Db      13 RVRVVRVRVRR 24
```

```
RESULT 6
US-09-785-058-5
; Sequence 5, Application US/09785058
; Publication No. US20030036627A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A 34001 / 072396.0222
; CURRENT APPLICATION NUMBER: US/09/785,058
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-058-5
```

```
Query Match          100.0%; Score 54; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 RVRVVRVRVRR 12
        |||||
Db      13 RVRVVRVRVRR 24
```

```
RESULT 7
US-09-785-059-6
; Sequence 6, Application US/09785059
; Patent No. US20020169279A1
; GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mietzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 12:01:36 ; Search time 6.38298 Seconds
(without alignments)
194.092 Million cell updates/sec

Title: US-10-079-075-4
Perfect score: 54
Sequence: 1 RVRVRVRRVRR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 segs, 103240269 residues
Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubppa/PCR_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubppa/PCRUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	12	US-09-785-059-4	Sequence 4, App11
2	54	100.0	12	US-10-079-075-4	Sequence 4, App11
3	54	100.0	12	US-09-785-058-4	Sequence 4, App11
4	54	100.0	24	US-09-785-059-5	Sequence 5, App11
5	54	100.0	24	US-10-079-075-5	Sequence 5, App11
6	54	100.0	24	US-09-785-058-5	Sequence 5, App11
7	54	100.0	36	US-09-785-059-6	Sequence 6, App11
8	54	100.0	36	US-10-079-075-6	Sequence 6, App11
9	54	100.0	36	US-09-785-058-6	Sequence 6, App11
10	54	100.0	42	US-09-785-059-7	Sequence 7, App11
11	54	100.0	42	US-10-079-075-7	Sequence 7, App11
12	54	100.0	42	US-09-785-058-7	Sequence 7, App11
13	54	100.0	48	US-09-785-059-8	Sequence 8, App11
14	54	100.0	48	US-10-079-075-8	Sequence 8, App11
15	54	100.0	48	US-09-785-058-8	Sequence 8, App11
16	47	87.0	12	US-09-785-059-9	Sequence 9, App11
17	47	87.0	12	US-10-079-075-9	Sequence 9, App11
18	47	87.0	12	US-09-785-058-9	Sequence 9, App11
19	47	87.0	24	US-09-785-059-10	Sequence 10, App11

20	47	87.0	24	9	US-10-079-075-10	Sequence 10, App1
21	47	87.0	24	9	US-09-785-058-10	Sequence 10, App1
22	47	87.0	36	9	US-09-785-059-11	Sequence 11, App1
23	47	87.0	36	9	US-10-079-075-11	Sequence 11, App1
24	47	87.0	36	9	US-09-785-058-11	Sequence 11, App1
25	47	87.0	48	9	US-09-785-059-12	Sequence 12, App1
26	47	87.0	48	9	US-10-079-075-12	Sequence 12, App1
27	47	87.0	48	9	US-09-785-058-12	Sequence 12, App1
28	35	64.8	28	9	US-09-785-059-1	Sequence 1, App11
29	35	64.8	28	9	US-10-079-075-1	Sequence 1, App11
30	35	64.8	28	9	US-09-785-058-1	Sequence 1, App11
31	35	64.8	31	9	US-09-785-059-2	Sequence 2, App11
32	35	64.8	31	9	US-10-079-075-2	Sequence 2, App11
33	35	64.8	31	9	US-09-785-058-2	Sequence 2, App11
34	35	63.0	147	9	US-10-137-765-60	Sequence 60, App1
35	35	63.0	147	9	US-10-146-337-60	Sequence 60, App1
36	34	63.0	485	10	US-09-808-483-12	Sequence 12, App1
37	34	63.0	535	10	US-09-808-483-10	Sequence 10, App1
38	33	61.1	868	9	US-09-938-406-1	Sequence 1, App11
39	33	61.1	2462	9	US-09-819-104A-5	Sequence 5, App11
40	32	59.3	48	9	US-10-083-357-1130	Sequence 1130, App
41	32	59.3	88	9	US-09-738-626-5561	Sequence 5561, App
42	32	59.3	226	9	US-09-860-670-139	Sequence 139, App
43	32	59.3	323	9	US-09-804-291-163	Sequence 163, App
44	32	59.3	323	10	US-09-886-035-163	Sequence 163, App
45	32	59.3	348	9	US-09-804-291-165	Sequence 165, App

ALIGNMENTS

RESULT 1
US-09-785-059-4
; Sequence 4, Application US/09785059
; Patent No. US20020169279A1
GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A33577 / 072396.0217
; CURRENT APPLICATION NUMBER: US/09785,059
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide derived from HIV-1
US-09-785-059-4
Query Match 100.0%; Score 54; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00092;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RVRVRVRRVRR 12
Db 1 RVRVRVRRVRR 12
RESULT 2
US-10-079-075-4
; Sequence 4, Application US/10079075
; Publication No. US20020188102A1
GENERAL INFORMATION:
; APPLICANT: Ronald C. Montelaro
; APPLICANT: Timothy A. Mielzner
; TITLE OF INVENTION: VIRUS DERIVED ANTIMICROBIAL PEPTIDES
; FILE REFERENCE: A34001-A / 072396.0222
; CURRENT APPLICATION NUMBER: US/10/079,075
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 12

PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

XX Example 1; SEQ ID No 3309; 1069pp; English.

XX
PS
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 82 AA;

Query Match 63.0%; Score 34; DB 22; Length 82;

Best Local Similarity 66.7%; Pred. No. 78;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVRVRRVRR 12

Db 4 RVAQFMRVRR 15

RESULT 15

ID AAU43929 standard; Protein; 165 AA.

XX AAU43929;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #4825.

XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/7L.

DR N-PSDB; AASS9521.

XX

PT

PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

XX Example 1; SEQ ID No 5124; 1069pp; English.

XX
PS
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 165 AA;

Query Match 63.0%; Score 34; DB 22; Length 165;

Best Local Similarity 70.0%; Pred. No. 1.6e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 11

Db 37 VTRVRRVRR 46

Search completed: June 9, 2003, 11:55:31
Job time : 13.9362 secs

Propionibacterium acnes polypeptides and nucleic acids useful for

PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 66.7%; Score 36; DB 21; Length 422;
Best Local Similarity 45.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 12
DB 25 IIRYRQIIR 35

RESULT 13
AAU49478
ID AAU49478 standard; Protein; 66 AA.

AC AAU49478;
XX
XX
DT 27-FEB-2002 (first entry)
XX

DE Propionibacterium acnes immunogenic protein #10374.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.
DR N-PSDB; AAS59545.

PT Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for
treating acne vulgaris -
XX
XX Example 1; SEQ ID No 10673; 10699p; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 66 AA;

Query Match 63.0%; Score 34; DB 22; Length 66;
Best Local Similarity 58.3%; Pred. No. 63;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVRVRRVRR 12
DB 1 RIRIVRRVRR 12

RESULT 14
AAU42114
ID AAU42114 standard; Protein; 82 AA.

AC AAU42114;
XX
XX
DT 27-FEB-2002 (first entry)
XX

DE Propionibacterium acnes immunogenic protein #3010.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.
DR N-PSDB; AAS59516.

PT Propionibacterium acnes polypeptides and nucleic acids useful for

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 26-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137282.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139442.
PR 17-JUN-1999; 99US-0139443.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142970.
PR 13-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143524.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144684.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145216.
PR	27-JUL-1999;	99US-0145293.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149930.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152353.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	23-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156556.

PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR		

XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Dramanac RT, Liu C, Tang YT;
DR WPI; 2001-639362/73.
DR N-PSDB; AAS92852.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 20; SEQ ID No 59024; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 807 AA;
SO

Query Match 68.5%; Score 37; DB 22; Length 807;
Best Local Similarity 72.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RVRVRVRVR 11
Db 511 RVTQVRRVRR 521

RESULT 11
AAG58025
ID AAG58025 standard; Protein; 117 AA.
XX
XX AAG58025;
AC
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 74852.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EPI033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135529.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140821.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

XX 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US08631.
XX
PF 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
PR
XX (HYSE-) HYSEQ INC.
PA
XX
PI Drmanac RT, Liu C, Tang YT;
DR WPI; 2001-639362/73.
DR N-PSDB; AAS92987.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
PS Claim 20; SEQ ID No 59159; 103pp; English.
XX
XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 804 AA;
Query Match 68.5%; Score 37; DB 22; Length 804;
Best Local Similarity 72.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 RVRVVRVRVR 11
Db 508 RVTQVVRVRVAR 518
RESULT 9
ABG29023
ID ABG29023 standard; Protein: 804 AA.
XX
AC ABG29023;
XX
DT 18-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #29014.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX OS
XX PN WO200175067-A2.
XX PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.
PF 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
PR
XX (HYSE-) HYSEQ INC.
PA
XX
PI Drmanac RT, Liu C, Tang YT;
DR WPI; 2001-639362/73.
DR N-PSDB; AAS93210.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
PS Claim 20; SEQ ID No 59382; 103pp; English.
XX
XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 804 AA;
Query Match 68.5%; Score 37; DB 22; Length 804;
Best Local Similarity 72.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 RVRVVRVRVR 11
Db 508 RVTQVVRVRVAR 518
RESULT 10
ABG28665
ID ABG28665 standard; Protein: 807 AA.
XX
AC ABG28665;
XX
DT 18-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #28656.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX OS
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.

Query Match 74.1%; Score 40; DB 21; Length 141;
 Best Local Similarity 81.8%; Pred. No. 14;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 12
 |||||:
 DB 5 VVRVRRLSRR 15

RESULT 6
 AAW37363
 ID AAW37363 standard; Protein; 320 AA.

XX AAW37363;
 AC
 XX
 DT 27-APR-1998 (first entry)
 XX
 DE PdbK (WbpK) protein involved in O-antigen synthesis and assembly.
 XX
 KM O antigen; PdbK; WbpK; lipopolysaccharide; infection; diagnosis;
 XX antibody; dehydratase.
 XX
 OS Pseudomonas aeruginosa PA01.
 XX
 PN MO9741234-A2.
 XX
 PD 06-NOV-1997.
 XX
 PF 30-APR-1997; 97WO-CA00295.
 XX
 PR 27-FEB-1997; 97US-0039473.
 XX
 PR 30-APR-1996; 96US-0016510.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Burrows L, Charter D, De Kievit T, Lam JS;
 XX
 DR WPI; 1997-549736/50.
 DR N-PSDB; AAT97221.
 XX
 PT Pseudomonas aeruginosa B-band lipopolysaccharide gene cluster -
 used for diagnosis of P. aeruginosa infection
 XX
 PS Claim 8; Page 123; 195pp; English.
 XX
 XX

This sequence comprises PdbK (WbpK), a Pseudomonas aeruginosa PA01 protein with dehydratase activity. Wzz (Pol), Peba (Wpba), PabB (WbpB), PabC (WbpC), PabD (WbpD), PabE (WbpE), Rfc (Wzy), PabF (WbpF), PabG (WbpG), PabH (WbpH), PabI (Wpbi), PabJ (WbpJ), PabK (WbpK), PabM (WbpM) and PabN (WbpN) (see AAW37349-56 and AAW37357-65, respectively) are claimed. They are involved in the synthesis and assembly of B-band lipopolysaccharide (i.e. O-antigen). Also claimed are: Wvrb (AAW37366) involved in ultraviolet repair and Hish and Hif (AAW37357-58) involved in histidine synthesis. All these proteins are encoded by the B-band gene cluster (see AAT97221) of P. aeruginosa PA01. Purified proteins can be obtained from CC transformed host cells and used to raise monoclonal or polyclonal antibodies. Such antibodies specifically recognise the B-band lipopolysaccharide and can be used in a claimed method for detecting P. aeruginosa in a sample, i.e. to diagnose infection.

XX
 SQ Sequence 320 AA;

Query Match 68.5%; Score 37; DB 18; Length 320;
 Best Local Similarity 72.7%; Pred. No. 99;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 12
 |||||:
 DB 28 VIAVRRVRR 38

RESULT 7

AAB66450
 ID AAB66450 standard; Protein; 331 AA.
 XX
 AC AAB66450;
 XX
 DT 06-APR-2001 (first entry)
 XX
 DE Pseudomonas aeruginosa WbpK06 carrying N-terminal hexahistidine tag.
 XX
 KM Pseudomonas aeruginosa; WbpK06; WbpM; UDP-N-acetylglucosamine;
 XX GalNAc; UDP-N-acetylglucosamine; GlcNAc; epimerase; WbpM assay.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN CA2305716-A1.
 XX
 PD 28-NOV-2000.
 XX
 PF 09-MAY-2000; 2000CA-2305716.
 XX
 PR 28-MAY-1999; 99US-0136564.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Creuzenet C, Burrows LL, Lam JS;
 XX
 DR WPI; 2001-169230/18.
 DR N-PSDB; AAF29636.
 XX
 PT Assaying for WbpM or its homologues in a sample comprises incubating
 sample with a substrate containing UDP-N-acetylglucosamine and/or
 PT UDP-N-acetylglucosamine where an increase in UDP-N-acetylglucosamine
 PT indicates the presence of WbpM -
 XX
 XX

PS Example 5; Fig 17; 63pp; English.

XX The present sequence is given in a specification relating to a method
 CC of assaying for WbpM or its homologue in a sample. The method comprises
 CC incubating a sample suspected of containing WbpM or its homologue and a
 CC substrate containing UDP-N-acetylglucosamine (GalNAc) and/or
 CC UDP-N-acetylglucosamine (GlcNAc), and quantifying the amount of
 CC UDP-GalNAc and UDP-GlcNAc in the sample after reaction, where an
 CC increase in UDP-GlcNAc reflects the presence of WbpM or its homologue.
 CC The assay can also be used to screen for inhibitors of WbpM. Antibodies
 CC can be used to detect WbpM or its homologues in a sample.
 XX

SQ Sequence 331 AA;

Query Match 68.5%; Score 37; DB 22; Length 331;
 Best Local Similarity 72.7%; Pred. No. 16+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 12
 |||||:
 DB 39 VIAVRRVRR 49

RESULT 8
 ABG28800
 ID ABG28800 standard; Protein; 804 AA.

XX
 AC ABG28800;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #28791.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 PN WO200175067-A2.

PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147263.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148655.
PR 17-JUN-1999;	99US-0139482.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149829.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0140353.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140354.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140355.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140685.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 02-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0155479.
PR 06-JUL-1999;	99US-0142390.	PR 22-SEP-1999;	99US-0155439.
PR 08-JUL-1999;	99US-0142803.	PR 23-SEP-1999;	99US-0155486.
PR 09-JUL-1999;	99US-0142920.	PR 24-SEP-1999;	99US-0155659.
PR 12-JUL-1999;	99US-0142977.	PR 28-SEP-1999;	99US-0156458.
PR 13-JUL-1999;	99US-0143542.	PR 29-SEP-1999;	99US-0156596.
PR 14-JUL-1999;	99US-0143624.	PR 04-OCT-1999;	99US-0157117.
PR 15-JUL-1999;	99US-0144005.	PR 05-OCT-1999;	99US-0157753.
PR 16-JUL-1999;	99US-0144085.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144086.	PR 07-OCT-1999;	99US-0158029.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145112.	PR 21-OCT-1999;	99US-0160814.
PR 23-JUL-1999;	99US-0145115.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160981.
PR 26-JUL-1999;	99US-0145276.	PR 23-OCT-1999;	99US-0160989.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145919.	PR 25-OCT-1999;	99US-0161406.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161359.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161361.
PR 03-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161920.
PR 04-AUG-1999;	99US-0147204.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147302.	PR 29-OCT-1999;	99US-0162142.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149992.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 74.1%; Score 40; DB 21; Length 48;
Best Local Similarity 81.8%; Pred. No. 4.5;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 5 VVRVRRVRR 12
5 VVRVRRLSRR 15

RESULT 5
ID AAG09943 standard; Protein; 141 AA.
AC AAG09943;
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 8071.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134372.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 74.1%; Score 40; DB 21; Length 48;
Best Local Similarity 81.8%; Pred. No. 4.5;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRVRR 12
|||:|
Db 5 VVRVRRLSRR 15

RESULT 4
AAG36558
ID AAG36558 standard; Protein; 48 AA.
XX
AC AAG36558;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44819.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135523.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143277.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135114.
PR 21-MAY-1999; 99US-0135333.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147203.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

XX Example 102; Page 84; 119pp; English.

PS The invention provides an antimicrobial compound (I) which is a peptide

XX having 8-50 amino acids, a net charge of 4, a hydrophobic moment (microh)

CC as a beta sheet which is 0.2 higher than its micro H as alpha helix, and

CC having detectable membrane disrupting activity against a microbial

CC pathogen, and substantially no membrane disrupting activity against

CC mammalian cells. (I) is useful for inhibiting microbial activity. (II)

CC has a detectable membrane disrupting activity against a pathogen. (II)

CC is useful for inhibiting non-microbial pathogenic activity also. (I) is also

CC useful for killing human sperm. The peptides are also provided in the

CC form of an expression vector comprising a nucleic acid encoding the

CC peptides. The peptides are useful for inhibiting the activity of

CC bacteria, and other microbial pathogens such as algae, fungi or protozoa

CC and for inhibiting non-microbial pathogens such as worms or arthropods,

CC and as spermicides for humans as the sperm membrane is atypical of human

CC cell membranes. (I) also has diagnostic uses e.g., in localizing an

CC infection or detecting sepsis. The peptides may act as binding molecules

CC and are useful to purify a target from blood, for qualitative or

CC quantitative analysis of analytes in in vitro sample, and for in vivo

CC imaging. Also, they are useful as molecular weight markers, as nutrient

CC source, as growth medium component for culturing microorganisms, as well

CC as a food ingredient for human consumption. The peptides have a greater

CC selectivity for bacterial versus mammalian lipids as compared to the

CC alpha helical peptides. Sequences AAG65536-47 represent amino acid

CC sequences of antimicrobial peptides.

SQ Sequence 18 AA;

XX

Query Match 88.9%; Score 48; DB 22; Length 18;

Best Local Similarity 50.0%; Pred. No. 0.075;

Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 RVRVVRVRVR 12

DB 6 RIRIRIRIRIR 17

RESULT 2

AAW05116

ID AAW05116 standard; peptide; 17 AA.

XX AAW05116;

AC

XX 19-JUN-1997 (first entry)

DT

XX Porcine somatotropin mimic.

DE

XX Porcine somatotropin mimic.

XX

KM porcine somatotropin; PST; growth promoter; helical conformation.

KX

XX Synthetic.

OS

XX WO9630405-A1.

PN

XX 03-OCT-1996.

PD

XX 15-MAR-1996; 96WO-US03490.

PF

XX 31-MAR-1995; 95US-0415239.

PR

XX (AMCY) AMERICAN CYANAMID CO.

PA

XX Buckwalter BL, Shieh H, Wang BS;

PI

XX WPI; 1996-485447/48.

DR

XX Peptide(s) mimicking a helical region of porcine somatotropin - used

PT in compositions to promote mammalian growth

PR

XX Claim 3; Page 17; 63pp; English.

PS

XX New peptides are disclosed which, by virtue of having certain

CC defined amino acids at every third or fourth residue, have a well

CC defined secondary structure which mimics the helical conformation

CC of a corresponding region of porcine somatotropin (PST). The peptides

CC enhance the activity of PST and promote the growth of warm-blooded

CC animals, especially pigs. They compete with PST for binding to the

CC PS-7.6 monoclonal antibody. The peptides have the generic sequence

CC XXXLXXIXXXVXX (I) or XXXXXLXXIXXXV (II); where residues X are

CC undefined other than the statement that the sequences differ from the

CC native sequence of PST. Formula (II) represents a peptide in which the

CC location of the essential amino acids is shifted by three amino acids,

CC representing almost one turn along the helix. Preferably X(2) of (II) is

CC Ile. Preferably the peptides contain Ser (as a promoter of helical

CC conformation) as the amino acid immediately amino-terminal to the first

CC Leu in (I) or to the first Ile of (II). Also, one or more of the first

CC or second Leu or the Val of (I) may be replaced by Nle. Furthermore, a

CC Cys residue may be added to either or both ends of the peptides.

CC The present sequence represents a specific example of the new

CC peptides.

CC

SQ Sequence 17 AA;

XX

Query Match 75.9%; Score 41; DB 17; Length 17;

Best Local Similarity 58.3%; Pred. No. 1.1;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVRVVRVRVR 12

DB 5 RIRIRIRIRIR 16

RESULT 3

AAG13632

ID AAG13632 standard; Protein; 48 AA.

XX AAG13632;

AC

XX 17-OCT-2000 (first entry)

DT

XX Arabidopsis thaliana protein fragment SEQ ID NO: 13198.

DE

XX

KM Protein identification; signal transduction pathway; metabolic pathway;

KW hybridization assay; genetic mapping; gene expression control; promoter;

KX termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

PD

XX 06-SEP-2000.

PF

XX 25-FEB-2000; 2000EP-0301439.

PR

XX 25-FEB-1999; 99US-0121825.

PR

XX 05-MAR-1999; 99US-0123180.

PR

XX 09-MAR-1999; 99US-0123548.

PR

XX 23-MAR-1999; 99US-0125788.

PR

XX 25-MAR-1999; 99US-0126264.

PR

XX 28-MAR-1999; 99US-0126785.

PR

XX 01-APR-1999; 99US-0127462.

PR

XX 06-APR-1999; 99US-0128234.

PR

XX 08-APR-1999; 99US-0128714.

PR

XX 16-APR-1999; 99US-0129845.

PR

XX 19-APR-1999; 99US-0130077.

PR

XX 21-APR-1999; 99US-0130449.

PR

XX 23-APR-1999; 99US-0130510.

PR

XX 28-APR-1999; 99US-0130891.

PR

XX 30-APR-1999; 99US-0131449.

PR

XX 30-APR-1999; 99US-0132048.

PR

XX 04-MAY-1999; 99US-0132484.

PR

XX 05-MAY-1999; 99US-0132485.

PR

XX 06-MAY-1999; 99US-0132486.

PR

XX 06-MAY-1999; 99US-0132487.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:35:06 ; Search time 12.9362 Seconds
(without alignments)
123.607 Million cell updates/sec

Title: US-10-079-075-4

Perfect score: 54

Sequence: 1 RVRRVRRVRR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	88.9	18	22	AA65539
2	41	75.9	17	17	AAW05116
3	40	74.1	48	21	AA613632
4	40	74.1	48	21	AA636558
5	37	68.5	141	21	AA609943
6	37	68.5	320	18	AAW37363
7	37	68.5	331	22	AA666450
8	37	68.5	804	22	ABG28800
9	37	68.5	804	22	ABG29023
10	37	68.5	807	22	ABG28665

11	36	66.7	117	21	AA658025	Arabidopsis thalia
12	36	66.7	422	21	AA637880	Arabidopsis thalia
13	34	63.0	66	22	AAU49478	Propionibacterium
14	34	63.0	82	22	AAU42114	Propionibacterium
15	34	63.0	165	22	AAU43929	Propionibacterium
16	34	63.0	178	22	AAU59336	Propionibacterium
17	34	63.0	209	21	AA603844	Human secreted pro
18	34	63.0	239	22	AB611801	Human secreted pro
19	34	63.0	485	22	AB663260	D. melanogaster 1i
20	34	63.0	485	22	AA680006	D. melanogaster 1i
21	34	63.0	485	22	AA675566	Amino acid sequenc
22	34	63.0	535	22	AA680005	D. melanogaster 1i
23	33	61.1	17	19	AAW47639	Antimicrobial pept
24	33	61.1	17	20	AAV32575	Antimicrobial pept
25	33	61.1	129	22	AAU56679	Propionibacterium
26	33	61.1	196	22	AB668506	Drosophila melano
27	33	61.1	201	22	ABG17909	Novel human diagno
28	33	61.1	308	21	AA614576	Arabidopsis thalia
29	33	61.1	308	21	AA649113	Arabidopsis thalia
30	33	61.1	309	21	AA614575	Arabidopsis thalia
31	33	61.1	309	21	AA649112	Arabidopsis thalia
32	33	61.1	316	21	AA614574	Arabidopsis thalia
33	33	61.1	316	21	AA649111	Arabidopsis thalia
34	33	61.1	316	23	AB693866	Herbicideally activ
35	33	61.1	337	21	AA626520	Arabidopsis thalia
36	33	61.1	576	22	AB664915	Drosophila melano
37	33	61.1	588	22	AA696513	Purative P. abyssi
38	33	61.1	844	19	AAW43073	HIV-1 gp120 protei
39	33	61.1	853	19	AAW43072	HIV-1 gp120 protei
40	33	61.1	865	19	AAW43072	HIV-1 gp120 protei
41	33	61.1	896	22	AB628532	Novel human diagno
42	33	61.1	1220	22	ABG24318	Novel human diagno
43	33	61.1	2675	21	AA607564	Protein encoded by
44	33	59.3	15	18	AAW33416	Leu,Arg analogue o
45	32	59.3	20	23	AA616414	ppG33 cationic pe

ALIGNMENTS

RESULT 1
ID AAG65539 standard; peptide, 18 AA.
XX AAG65539;
AC
XX
XX
XX 30-NOV-2001 (first entry)
DE Peptide sequence used in the course of the invention.
XX
XX Antimicrobial; microbial membrane disrupter; gene therapy; pathogen;
XX
XX Antimicrobial; microbial membrane disrupter; gene therapy; pathogen;
XX
XX Antimicrobial; microbial membrane disrupter; gene therapy; pathogen;
XX
XX Synthetic.
XX
XX WO200160162-A2.
XX
XX 23-AUG-2001.
XX
XX 15-FEB-2001; 2001WO-US04822.
XX
XX 15-FEB-2000; 2000US-0182495.
XX
XX (UYOH-) UNIV OHIO.
XX
XX Blazyk JF;
XX
XX WPI; 2001-565322/63.
XX
XX Novel peptides having antimicrobial activity have positive charge to
XX selectively disrupt microbial membranes, assume beta sheet structure in
XX membrane environment and are substantially amphipathic in beta sheet
XX structure

SQ SEQUENCE 364 AA; 41674 MW; 2F1B38SC3722D3D CRC64;

Query Match 66.7%; Score 36; DB 5; Length 364;

Best Local Similarity 72.7%; Pred. NO. 1.6e+02; Mismatches 0; Gaps 0;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVRVVRVVR 11
| | | | |
| | | | |

Db 74 REVVARBMVR 84

Search completed: June 9, 2003, 12:01:03
Job time : 13.0213 secs

RL ProC. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010420; AAM02504.1; -.
 KW Complete proteome.
 SQ SEQUENCE 214 AA; 25302 MW; 3812F54C345A0412 CRC64;
 Query Match 66.7%; Score 36; DB 17; Length 214;
 Best Local Similarity 54.5%; Pred. No. 93;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RVVVRVRRVRR 12
 DB 19 RVLRVLRVRR 30
 RESULT 13
 ID 08U603 PRELIMINARY; PRT; 221 AA.
 AC 08U603;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ABC transporter, membrane spanning protein.
 GN ATU4754 OR AGR L 252.
 OS Agrobacterium tumefaciens (strain CS8 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; Pubmed=11743193;
 RA Wood D.W., Serubal J.C., Kaul R., Monke D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
 RA Chapman P., Clendinning J., Decherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Tung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley G., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 CS8";
 RT Science 294:2317-2323(2001).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=21608551; Pubmed=11743194;
 RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Currollo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
 RA Houtmel K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
 RA Flanagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens CS8";
 RT Science 294:2323-2328(2001).
 DR EMBL; AE009404; AAL45548.1; -.
 DR EMBL; AE008210; AAK8696.1; -.
 KW Complete proteome.
 SQ SEQUENCE 221 AA; 24149 MW; D5F58B6910660BC CRC64;
 Query Match 66.7%; Score 36; DB 16; Length 221;
 Best Local Similarity 54.5%; Pred. No. 96;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 VVRVRRVRR 12
 DB 201 IIRVIRVRR 211
 RESULT 14
 ID 08TW69 PRELIMINARY; PRT; 241 AA.
 AC 08TW69;
 KW

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Uncharacterized protein.
 GN MK167.
 OS Methanopyrus kandleri;
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; Pubmed=11930014;
 RA Stesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natile D.A., Rogozin I.B., Tatunov R.L., Wolf Y.I., Steetler K.O.,
 RA Mal'kh A.G., Koonin E.V., Kozaykin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010407; AAM02380.1; -.
 KW Complete proteome.
 SQ SEQUENCE 241 AA; 25177 MW; 3A8ACEF5874735E1 CRC64;
 Query Match 66.7%; Score 36; DB 17; Length 241;
 Best Local Similarity 54.5%; Pred. No. 1e+02;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VVRVRRVRR 12
 DB 188 IIRVIRVRR 198
 RESULT 15
 ID 017983 PRELIMINARY; PRT; 364 AA.
 AC 017983;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 41.7 kDa protein.
 GN C14E2.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; Pubmed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium";
 RT Science 282:2012-2018(1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=BRISTOL N2;
 RC Bentley D., Gattung S.;
 RT "The sequence of C. elegans cosmid C14E2";
 RT Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission";
 RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; U46671; AAB5747.1; -.
 DR HSSP; P00503; IAS.
 DR InterPro; IPR004839; AminoTransf1/2.
 DR InterPro; IPR000796; Aspartate_sub.
 DR Pfam; PF00155; aminotran_1_2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 KW Hypothetical protein.

Db 813 RVIEVRRIR 823

RESULT 9
Q9E1S5 PRELIMINARY; PRT; 856 AA.

AC Q9E1S5; 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Envelope protein (fragment).
GN ENV.

OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;

RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=20442410; PubMed=10984542;
RA Cleghorn F.R., Jack N., Carr J.K., Edwards J., Mahabir B., Sill A.,
McDaniel C.B., Connolly S.M., Goodman D., Bennett R.O., O'Brien T.R.,
RT "A distinctive clade B HIV type 1 is heterosexually transmitted in
Trinidad and Tobago."
RL Proc. Natl. Acad. Sci. U.S.A. 97:10532-10537(2000).

DR EMBL; AF277060; AG22505.1; -
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.

KW Transmembrane.
FT NON TER 856

SQ SEQUENCE 856 AA; 96999 MW; 7C0D5F88F5C28915 CRC64;

Query Match 68.5%; Score 37; DB 15; Length 856;
Best local similarity 63.6%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVVRRVRRV 11
DB 828 RVIEVRRIR 838

RESULT 10
Q98LM6 PRELIMINARY; PRT; 1015 AA.

AC Q98LM6; 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable conjugugal transfer protein, TraA.

GN M10964.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,
Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).

DR EMBL; AP002996; BAB48437.1; -
DR InterPro: IPR005053; MoB_Mobl.
DR Pfam; PF03389; MoB_Mobl; 1.
KW Complete proteome.

SQ SEQUENCE 1015 AA; 112159 MW; 64C2B1B74280D9A9 CRC64;

Query Match 68.5%; Score 37; DB 16; Length 1015;
Best local similarity 70.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVVRRVRRV 10
DB 753 RVIEVRRIR 762

RESULT 11
Q9KXQ3 PRELIMINARY; PRT; 121 AA.

AC Q9KXQ3; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC01497 OR SC9C5.21C.

OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetes.

OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornby T., Howarth S.,
Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajendram M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)."
RL Nature 417:141-147(2002).

DR EMBL; AL357523; CAB93377.1; -
SQ SEQUENCE 121 AA; 12888 MW; AF2BE07727DB24D1 CRC64;

Query Match 66.7%; Score 36; DB 16; Length 121;
Best local similarity 72.7%; Pred. No. 53;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRRVRRV 12
DB 43 VVRRVRRV 53

RESULT 12
Q8TVU7 PRELIMINARY; PRT; 214 AA.

AC Q8TVU7; 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Uncharacterized protein specific for M.kandleri, MK-1 family.

GN MK1291.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.

OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=AV19 / DSM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
Nichezhina O.V., Shakhova V.V., Belova G.I., Aravind L.,
Natale D.A., Koonin E.V., Kozlov R.L., Wolf Y.I., Stetter K.O.,
RA Malyn A.G., Koonin E.V., Kozlov R.L., Wolf Y.I., Stetter K.O.,
RT "The complete genome of hyperthermophilic Methanopyrus kandleri AV19
and monophyly of archaeal methanogens."

RT "The complete genome of hyperthermophilic Methanopyrus kandleri AV19
and monophyly of archaeal methanogens."

DE Hypothetical protein DR0533.
GN DR0533.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Moffet K.S., Qin H., Jiang L., Pamphile W., Crosby W., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Usterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AF001911; AAF10115.1; -.
DR TIGR; DR0533; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 276 AA; 30266 MW; 533F34C6E434D94A CRC64;

Query Match 68.5%; Score 37; DB 16; Length 276;
Best Local Similarity 66.7%; Pred. No. 80;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVVRVRRVRR 12
DB 229 RVVRVRRVRR 240

RESULT 6
Q9H281 PRELIMINARY; PRT; 316 AA.
ID Q9H281;
AC Q9H281;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Probable NAD-dependent epimerase/dehydratase WbpK.
GN WbpK OR PA3146.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RA MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004738; AAG06534.1; -.
KW Complete proteome.
SQ SEQUENCE 316 AA; 33989 MW; B61A722C82501DBD CRC64;

Query Match 68.5%; Score 37; DB 16; Length 316;
Best Local Similarity 72.7%; Pred. No. 92;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVVRVRRVRR 12
DB 24 VVVRVRRVRR 34

RESULT 7
ID P72144 PRELIMINARY; PRT; 320 AA.
AC P72144;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE WbpK.
GN WbpK.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01;
RA MEDLINE=97093969; PubMed=8939432;
RA Burrows L.L., Charter D.F., Lam J.S.;
RT "Molecular characterization of the Pseudomonas aeruginosa serotype O5
(PAO1) B-band lipopolysaccharide gene cluster.";
RL Mol. Microbiol. 22:481-495(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01;
RA Burrows L.L., Charter D.F., Lam J.S.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50396; AAC45865.1; -.
SQ SEQUENCE 320 AA; 34419 MW; 3AEC035FD3BFD40D CRC64;

Query Match 68.5%; Score 37; DB 2; Length 320;
Best Local Similarity 72.7%; Pred. No. 93;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VVVRVRRVRR 12
DB 28 VVVRVRRVRR 38

RESULT 8
Q9E1S4 PRELIMINARY; PRT; 841 AA.
ID Q9E1S4;
AC Q9E1S4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Envelope protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20442410; PubMed=10984542;
RA Cleghorn F.R., Jack N., Carr U.K., Edwards J., Mahabir B., Sill A.,
RA McDaniel C.B., Connolly S.M., Goodman D., Bennett R.O., O'Brien T.R.,
RA Weinhold K.J., Bartholomew C., Blattner W.A., Greenberg M.L.;
RT "A distinctive clade B HIV type 1 is heterosexually transmitted in
Trinidad and Tobago.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10532-10537(2000).
DR EMBL; AF277062; AAG22507.1; -.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 841
SQ SEQUENCE 841 AA; 95532 MW; 6A25B9FPAE901B7B CRC64;

Query Match 68.5%; Score 37; DB 15; Length 841;
Best Local Similarity 63.6%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVVRVRRVRR 11

QY 1 RVVVRVRRVRR 12
 ||:|:|:|:|:|
 DB 69 RIVKVRKVKR 80

RESULT 2

Q92YM6 PRELIMINARY; PRT; 561 AA.

AC 092YM6; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative methyl-accepting chemotaxis protein.
 GN RA0847 OR SMA1556.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=1021;
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 Bariloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
 Gurial M., Hong A., Hutzar L., Hyman R.W., Kahn D., Kahn M.L.,
 Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 Yan K.-C., Davis R.W., Federpiet N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 Sinorhizobium meliloti pSyma megaplasmid."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 DR EMBL: AE007271; AAK65505.1; -
 DR InterPro: IPR004089; ChmTaxis_transd.
 DR InterPro: IPR003660; HAMP.
 DR Pfam: PF00672; HAMP; 1.
 DR Pfam: PF00015; MCPsignal; 1.
 KM Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 561 AA; 59996 MW; 172840DC7D2AB413 CRC64;

Query Match 72.2%; Score 39; DB 16; Length 561;
 Best Local Similarity 72.7%; Pred. No. 73;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVVRVRRVRR 12
 ||:|:|:|:|:|
 DB 208 VLRVRLRR 218

RESULT 3

Q8YH48 PRELIMINARY; PRT; 279 AA.

AC 08YH48; 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Dihydropterate synthase (EC 2.5.1.15).

GN BMEI0956.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Ios T.,
 Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 Rajlowski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Lelesson J.-J.,
 Harelorn R., Kyriades N., Overbeek R.;
 RA "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009536; AAL52137.1; -
 DR InterPro: IPR000489; Dhdtcpt_synt.
 DR Pfam: PF00809; Pterin bind; 1.
 DR PROSITE: PS00792; DHS_1; UNKNOWN_1.
 DR PROSITE: PS00793; DHS_2; 1.
 KW Transferase; Complete proteome.
 SQ SEQUENCE 279 AA; 30255 MW; 8FECF6CED5A03B79 CRC64;

Query Match 70.4%; Score 38; DB 16; Length 279;
 Best Local Similarity 66.7%; Pred. No. 55;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVVVRVRRVRR 12
 ||:|:|:|:|:|
 DB 81 RVVVRVRRVRR 92

RESULT 4

Q74749 PRELIMINARY; PRT; 867 AA.

AC 074749; 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Envelope glycoprotein.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=96190564; PubMed=8627686;
 RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
 Karlsson G., Sodroski J., Morgado M., Galvao-Castro B., G.M.,
 von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
 Hahn B.H.;
 RT "Molecular cloning and analysis of functional envelope genes from
 human immunodeficiency virus type 1 sequence subtypes A through G. The
 WHO and NIAID Networks for HIV Isolation and Characterization."
 RL J. Virol. 70:1651-1667(1996).
 RN [2]
 RP SEQUENCE FROM N.A.

RA DAIDS Variation Program;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

RA MacInnes K.A.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U04908; AAB05049.1; -
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 867 AA; 98632 MW; F0780669D709D596 CRC64;

Query Match 70.4%; Score 38; DB 15; Length 867;
 Best Local Similarity 63.6%; Pred. No. 1.7e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVVVRVRRVRR 11
 ||:|:|:|:|:|
 DB 839 RIVVVRVRR 849

RESULT 5

Q9RWY2 PRELIMINARY; PRT; 276 AA.

AC 09RWY2; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:43:01 ; Search time 11.0213 Seconds

(without alignments)
224.345 Million cell updates/sec

Title: US-10-079-075-4

Perfect score: 54

Sequence: 1 RVRVRVRRVRR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMML_21.*
2: sp_archaea.*
3: sp_bacteria.*
4: sp_fungi.*
5: sp_human.*
6: sp_invertebrate.*
7: sp_mammal.*
8: sp_munc.*
9: sp_organelle.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virius.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	79.6	2351	10	O23372 arabidopsis
2	39	72.2	561	16	O92Ym6 rhizobium m
3	38	70.4	279	16	O8YH48 bruceella me
4	38	70.4	867	15	O74749 human immun
5	37	68.5	276	16	O9RWY2 deinococcus
6	37	68.5	316	16	O9H281 pseudomonas
7	37	68.5	320	2	P72144
8	37	68.5	841	15	O9E1S4 human immun
9	37	68.5	856	15	O9E1S5 human immun
10	37	68.5	1015	16	O9E1M6 rhizobium 1
11	36	66.7	121	16	O9KXQ3 streptomyce
12	36	66.7	214	17	O8TVU7 mechanopyru
13	36	66.7	221	16	O8U6O3 agrobacteri
14	36	66.7	241	17	O8TW69 mechanopyru
15	36	66.7	364	5	O17983 caenorhabdi
16	36	66.7	422	10	O9LUV7 arabidopsis

17	36	66.7	740	5	O21426	O21426 caenorhabdi
18	36	66.7	838	15	O9DVL4	O9DVL4 human immun
19	36	66.7	841	15	O41556	O41556 human immun
20	36	66.7	852	15	O69992	O69992 human immun
21	36	66.7	1677	11	O70373	O70373 mus musculu
22	35	64.8	87	2	O44954	O44954 bacillus br
23	35	64.8	117	17	O9V1Z0	O9V1Z0 pyrococcus
24	35	64.8	125	16	O99Q10	O99Q10 caulobacter
25	35	64.8	237	2	O67996	O67996 bruceella ab
26	35	64.8	247	12	O84507	O84507 paramacium
27	35	64.8	510	16	O9CD61	O9CD61 mycobacteri
28	35	64.8	732	10	O9XE45	O9XE45 arabidopsis
29	35	64.8	838	15	O8UTC7	O8UTC7 human immun
30	35	64.8	858	15	O87628	O87628 chimpanzee
31	35	64.8	862	15	O12004	O12004 chimpanzee
32	35	64.8	862	15	O12010	O12010 chimpanzee
33	35	64.8	868	15	O12009	O12009 chimpanzee
34	35	64.8	870	15	O12011	O12011 chimpanzee
35	35	64.8	874	15	O12003	O12003 chimpanzee
36	35	64.8	1015	16	O98B51	O98B51 rhizobium 1
37	35	64.8	1124	4	O9Y6S2	O9Y6S2 homo sapien
38	34	63.0	56	15	O56188	O56188 human immun
39	34	63.0	56	15	O56189	O56189 human immun
40	34	63.0	56	15	O56191	O56191 human immun
41	34	63.0	56	15	O56192	O56192 human immun
42	34	63.0	56	15	O56193	O56193 human immun
43	34	63.0	56	15	O56194	O56194 human immun
44	34	63.0	68	2	O9EVQ1	O9EVQ1 escherichia
45	34	63.0	153	6	O95L19	O95L19 sus scrofa

ALIGNMENTS

RESULT 1

ID O23372 PRELIMINARY: PRT; 2351 AA.

AC O23372;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein.
GN A74G15180.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Stiekema W., Murphy G., Wandutt R., Pohl T., Terryn N.,
RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
RA Puigdomenech P., Hatzopoulos P., Obermaier B., Duetschhoff A.,
RA Jones J., Palme K., Ansgorge W., Delseny M., Bancroft I., Mewes H.W.,
RA Schellier C., Chalwalz N.,
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
DR EMBL; Z97338; CAB10297.1; -
DR EMBL; AL161540; CAB78560.1; -
DR InterPro; IP001214; SET.
DR Pfam; PF00856; SET; 1.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS50280; SET; 1.
KW Hypothetical protein.
SQ SEQUENCE 2351 AA; 265519 MW; F403E3577C0F4C61 CRC64;

Query Match 79.6%; Score 43; DB 10; Length 2351;
Best Local Similarity 58.3%; Pred. No. 61;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS
CC BACTERIA.
CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M09410; AAA26762.1; ALT_INIT.
DR HSSP; P00579; ISIG.
DR InterPro; IPR000943; Sigma_70.
DR Pfam; PF00140; sigma70; 1.
DR PRINTS; PR00046; SIGMA70FCT.
DR PROSITE; PS00715; SIGMA70_1; 1.
DR PROSITE; PS00716; SIGMA70_2; 1.
KW Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
KW DNA-binding.
FT DOMAIN 108 121 POLYMERASE CORE BINDING (POTENTIAL).
FT DNA BIND 278 297 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 317 AA; 35616 MW; 62AED08CC61A6600 CRC64;

Query Match 61.1%; Score 33; DB 1; Length 317;
Best Local Similarity 63.6%; Pred. No. 55;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVRVVRVVR 11
Db 167 RVRVVRRLQ 177

RESULT 15
Y724 METH STANDARD; PRT; 336 AA.
AC O26820;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MTH724.
GN MTH724.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Baahirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadatoro R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Carnao A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- SIMILARITY: BELONGS TO THE UPF0020 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AE000852; AAB85229.1; -.
DR HSSP; P14385; 2ADM.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000241; RN4_methylase.
DR InterPro; IPR000051; SAM_bind.
DR InterPro; IPR004114; THUMP_dom.
DR Pfam; PF01170; UPF0020; 1.
DR TIGRPFAM; TIGR01177; TIGR01177; 1.
DR PROSITE; PS01261; UPF0020; 1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 336 AA; 38019 MW; D305B3FADF0E1F5CC CRC64;

Query Match 61.1%; Score 33; DB 1; Length 336;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RVRVVRVVR 9
Db 328 RVRVVRERV 336

Search completed: June 9, 2003, 11:56:36
Job time : 4.42553 secs.
```

```

CC or send an email to license@isb-sib.ch).
CC EMBL, D63796; BAA09865.1; -.
DR Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
SQ SEQUENCE 32 AA; 4220 MW; E10EC3A5B2ED6803 CRC64;

Query Match
Best Local Similarity 61.1%; Score 33; DB 1; Length 32;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RVVVRVVRVVR 12
DB 19 RVVRRRRVRRR 30

RESULT 12
RM16 PROMI STANDARD; PRT; 135 AA.
AC P46751;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Mitochondrial 60S ribosomal protein L16.
GN RPL16.
OS Prototheca wickerhamii.
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OX Chlorellaceae; Prototheca.
RX NCBI_TaxId=311;
RP SEQUENCE FROM N.A.
RC STRAIN=263-11;
RA MEDLINE=94180393; PubMed=8133522;
RT Wolff G., Plante I., Lang B.F., Kueck U., Burger G.;
"Complete sequence of the mitochondrial DNA of the chlorophyte alga
Prototheca wickerhamii. Gene content and genome organization.",
J. Mol. Biol. 237:75-86(1994).
RL J. Mol. Biol. 237:75-86(1994).
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U02970; AAD12638.1; -.
DR InterPro; IPR000114; Ribosomal L16.
DR Pfam; PF00252; Ribosomal L16; 1.
DR PRINTS; PRO0060; RIBOSOMAL16.
DR TIGRPFAM; TIGR01164; rplp_bact; 1.
DR PROSITE; PS00586; RIBOSOMAL_L16_1; 1.
DR PROSITE; PS00701; RIBOSOMAL_L16_2; 1.
KW Ribosomal protein; Mitochondrion.
SQ SEQUENCE 135 AA; 15311 MW; 1C240414BC5669E5 CRC64;

Query Match
Best Local Similarity 61.1%; Score 33; DB 1; Length 135;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RVVVRVVRVVR 12
DB 45 RVEAVRVRVTR 56

RESULT 13
Y118 VIRCH STANDARD; PRT; 314 AA.
ID Y118 VIRCH
AC 09K03;
RT 16-OCT-2001 (Rel. 40, Created)

```

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical adenine-specific methylase VC2118 (EC 2.1.1.72).
GN VC2118.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_TaxId=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Halt D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dighe I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC -1- SIMILARITY: BELONGS TO THE N6-METHYLTRANSFERASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AE004285; AAF95263.1; -.
DR TIGR; VC2118;
DR InterPro; IPR004556; HemK.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000051; SAM_bind.
DR TIGRPFAM; TIGR00536; hemk_fam; 1.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Hypothetical protein; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 314 AA; 35578 MW; BFSB9EA0FDEDBAF CRC64;

Query Match
Best Local Similarity 61.1%; Score 33; DB 1; Length 314;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVVVRVVR 12
DB 78 LRVVRVVR 87

RESULT 14
RPSD STRAU STANDARD; PRT; 317 AA.
ID RPSD STRAU
AC P27785;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RNA polymerase sigma factor rpoD.
GN RPOD.
OS Streptomyces aureofaciens.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=1894;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9308396; PubMed=1452038;
RA Kormanec J., Farkasovsky M., Potuckova L.;
RT "Four genes in Streptomyces aureofaciens containing a domain
RT characteristic of principal sigma factors.",
RT Gene 122:63-70(1992).

```

RC STRAIN=LT2 / SGGC1412 / ATCC 700720;
 RA MEDLINE=21534948; PubMed=1677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RA "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF COHYRIC ACID TO
 CC COBINAMIDE. ADDITION OF AMINOPROPANOL ON THE F CARBOXYLIC
 CC GROUP.
 CC -1- PATHWAY: Cobalamin biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE COB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L12006; AAA27253.1; -
 CC EMBL: AE008789; AAL20938.1; -
 CC STyGene: SGI0035; ChIB.
 CC DR InterPro: IPR004485; ChIB.
 CC DR Pfam: PF03186; COB_CbIB; 1.
 CC DR TIGR: TIGR00380; ChIB; 1.
 CC KM Cobalamin biosynthesis; Transmembrane; Complete proteome.
 CC FT TRANSMEM 3 23 POTENTIAL.
 CC FT TRANSMEM 56 76 POTENTIAL.
 CC FT TRANSMEM 82 102 POTENTIAL.
 CC FT TRANSMEM 153 173 POTENTIAL.
 CC FT TRANSMEM 296 316 POTENTIAL.
 CC SQ SEQUENCE 319 AA; 35370 MW; 540PDP93E0796214 CRC64;
 Query Match 63.0%; Score 34; DB 1; Length 319;
 Best Local Similarity 50.0%; Pred. No. 37;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RVRVVRVVR 12
 ID 1 RVRVVRVVR 12
 Db 31 RLTFVQRIVR 42
 RESULT 10
 YE40 ARCFU STANDARD; PRT; 613 AA.
 AC 028832;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1440.
 GN AF1440.
 OS Archaeoglobus fulgidus.
 OC Archaea, Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OC NCBI_TaxID=2234;
 OX RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klein H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gilm M., Hickey E.K., Peterson J.D.,
 RA Rischmann D.L., Kerlavage A.R., Graham D.B., Kyriakides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kitzman E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus S.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RA "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 CC -1- SIMILARITY: BELONGS TO THE GATB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE001004; AAB9807.1; -
 CC TIGR: AF1440; -
 CC DR InterPro: IPR004115; GAD_dom.
 CC DR InterPro: IPR003789; GATB_Ydeg.
 CC DR InterPro: IPR004414; GATB_Yel.
 CC DR InterPro: IPR001773; Gln_amlidtransfB.
 CC DR Pfam: PF01162; GATB; 1.
 CC DR Pfam: PF02637; DUF186; 1.
 CC DR Pfam: PF02934; GATB_N; 1.
 CC DR Pfam: PF02938; GAD; 1.
 CC DR TIGR: TIGR00134; GATB_Yel; 1.
 CC DR PROSITE: PS01234; GATB; 1.
 CC KM Hypothetical protein; Complete proteome.
 CC SQ SEQUENCE 613 AA; 69023 MW; 38B139182878529 CRC64;
 Query Match 63.0%; Score 34; DB 1; Length 613;
 Best Local Similarity 50.0%; Pred. No. 74;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 RVRVVRVVR 12
 ID 1 RVRVVRVVR 12
 Db 377 RVERALRIIR 388
 RESULT 11
 PRT_ORLYA STANDARD; PRT; 32 AA.
 AC 091185;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protamine.
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OC NCBI_TaxID=8090;
 OX RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Orange-red; TISSUE=Testis;
 RA Tamura M., Yamamoto H., Onitake K.;
 RA "Cloning of protamine cDNA of the medaka (*Oryzias latipes*) and its
 RT expression during spermatogenesis.";
 RL Dev. Growth Differ. 36:419-425(1994).
 CC -1- FUNCTION: PROTEIN SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Testis.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

entires requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AE000321; AAC5390.1; ALT_INIT.
 CC EMBL; D90864; BAA16188.1; ALT_INIT.
 DR EMBL; D90863; BAA16188.1; ALT_INIT.
 DR EMBL; M27714; NOT_ANNOTATED_CDS.
 DR EMBL; Y00720; NOT_ANNOTATED_CDS.
 DR EMBL; M3021; NOT_ANNOTATED_CDS.
 DR Ecogen; EG12449; YFCB.
 DR InterPro; IPR004556; HemK.
 DR InterPro; IPR002052; N6_Mtase.
 DR InterPro; IPR000051; SAM_bind.
 DR TIGR; TIGR00536; hemK_fam. 1.
 DR PROSITE; PS00092; N6_MTASE; 1.
 KW Hypothetical protein; Transferase; Methyltransferase;
 KW Complete proteome.
 SQ SEQUENCE 310 AA; 35001 MW; 3EB8F2D45AFC3760 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 310;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 RVRARVVR 12
 DB 75 RIVERVIR 83

RESULT 8
 YFCB_SALTY STANDARD; PRT; 310 AA.
 AC P39201;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical adenine-specific methylase yfcB (EC 2.1.1.72).
 GN YFCB OR STM2385 OR STY2617.
 OS Salmonella typhimurium, and
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxId=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stonking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RT Nature 413:852-856(2001).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Hungai K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Hogue A., Hien T.T., Holroyd S., Jajels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gea P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RA "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18.";
 RT Nature 413:848-853(2001).
 RL [3]
 RN SEQUENCE OF 225-310 FROM N.A.
 RC SPECIES=S.typhi; STRAIN=TY2;
 RX MEDLINE=90218018; PubMed=2182772;

RA Charles I.G., Lamb H.K., Pickard D., Dougan G., Hawkins A.R.;
 RT "Isolation, characterization and nucleotide sequences of the aroC
 RT genes encoding chorismate synthase from Salmonella typhi and
 RT Escherichia coli.";
 RL J. Gen. Microbiol. 136:353-358(1990).
 RN [4]
 RP IDENTIFICATION AND CONCEPTUAL TRANSLATION.
 RC SPECIES=S.typhi;
 RX MEDLINE=95075659; PubMed=7984428;
 RA Borodovsky M., Rudd K.E., Koonin E.V.;
 RT "Intrinsic and extrinsic approaches for detecting genes in a
 RT bacterial genome.";
 RL Nucleic Acids Res. 22:4766-4767(1994).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
 CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
 CC -1- SIMILARITY: BELONGS TO THE N6-METHYLTRANSFERASE FAMILY.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
 CC in positions 233 and 306.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE008807; AAL21286.1; -;
 CC DR EMBL; AL627274; CAD07617.1; -;
 DR EMBL; M27715; NOT_ANNOTATED_CDS.
 DR Scyene; SG7777; YFCB.
 DR InterPro; IPR004556; HemK.
 DR InterPro; IPR002052; N6_Mtase.
 DR InterPro; IPR000051; SAM_bind.
 DR TIGR; TIGR00536; hemK_fam. 1.
 DR PROSITE; PS00092; N6_MTASE; 1.
 KW Hypothetical protein; Transferase; Methyltransferase;
 KW Complete proteome.
 SQ SEQUENCE 310 AA; 35063 MW; 0205E269A8A28638 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 310;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 RVRARVVR 12
 DB 75 RIVERVIR 83

RESULT 9
 CBIB_SALTY STANDARD; PRT; 319 AA.
 AC C05600;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cbibi protein.
 GN CBIB OR STM2034.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxId=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RX MEDLINE=93273696; PubMed=8501034;
 RA Roth J.R., Lawrence J.G., Rubenfield M., Kieffer-Higgins S.,
 RA Church G.M.;
 RA "Characterization of the cobalamin (vitamin B12) biosynthetic genes
 RT of Salmonella typhimurium.";
 RT J. Bacteriol. 175:3303-3316(1993).
 RL [2]
 RN SEQUENCE FROM N.A.

```

DR MGD; MG1:98020; Rp119.
DR InterPro: IPR000136; Ribosomal_L19e.
DR Pfam: PF01280; Ribosomal_L19e; 1.
DR PROSITE: PS00526; RIBOSOMAL_L19E; 1.
KW Ribosomal protein.
FT CONFLICT 179 179 A -> S (IN REF. 4).
SQ SEQUENCE 196 AA; 23466 MW; 4AF506393E526216 CRC64;

Query Match
Best Local Similarity 63.0%; Score 34; DB 1; Length 196;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RVRVVRVVR 12
Db 97 RRRRLRLRLR 108

RESULT 6
URK_BACHD
ID URK_BACHD STANDARD; PRT; 211 AA.
AC 09KDD8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Uridine kinase (EC 2.7.1.48) (uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR BH1275.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AP001511; BAB04994.1; -.
DR InterPro: IPR001324; PRK.
DR InterPro: IPR000764; Uridine_kin.
DR Pfam: PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIBKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 211 AA; 24387 MW; CAAFA2CB0030520B CRC64;

Query Match
Best Local Similarity 55.6%; Score 34; DB 1; Length 211;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VRRVVRVVR 11
Db 137 IRIIRRVVR 145

```

```

RESULT 7
YFCB_ECOLI
ID YFCB_ECOLI STANDARD; PRT; 310 AA.
AC P39159; P76252; P76939;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical adenine-specific methylase yfcB (EC 2.1.1.72).
GN YFCB OR B2330.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1233-1238(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Sato Y., Sivasubram S.,
RA Tagami H., Takahashi T., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horikuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [3]
RP SEQUENCE OF 172-310 FROM N.A.
RX MEDLINE=90218018; PubMed=2182772;
RA Charles I.G., Lamb H.K., Pickard D., Dougan G., Hawkins A.R.;
RT "Isolation, characterization and nucleotide sequences of the aroC
RT genes encoding chorismate synthase from Salmonella typhi and
RT Escherichia coli".
RL J. Gen. Microbiol. 136:353-358(1990).
RN [4]
RP SEQUENCE OF 159-310 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88293429; PubMed=2969724;
RA White P.J., Millar G., Coggin J.R.;
RT "The overexpression, purification and complete amino acid sequence of
RT chorismate synthase from Escherichia coli K12 and its comparison with
RT the enzyme from Neurospora crassa".
RL Biochem. J. 251:313-322(1988).
RN [5]
RP IDENTIFICATION.
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Ruda K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
RT bacterial genome.";
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC adenosyl-L-homocysteine + DNA 6-methylaminouracil.
CC -1- SIMILARITY: BELONGS TO THE N6-METHYLTRANSFERASE FAMILY.
CC -1- CAUTION: REF. 4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
CC FRAMESHIFTS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```



```

CC CC VIRAL PROCHROMATIN FOR ENCAPSIDATION BY VIRTUE OF THE TWO
CC CC BASIC DOMAINS (BY SIMILARITY).
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U44123; AAB16758.1; -.
CC CC Core protein; DNA-binding; late protein.
CC CC PROPEP 1 26 BY SIMILARITY.
CC CC CHAIN 27 41 LATE L2 MU CORE PROTEIN.
CC CC PROPEP 42 70 BY SIMILARITY.
CC CC SITE 26 27 CLEAVAGE (BY ADENOVIRUS PROTEASE)
CC CC (POTENTIAL).
CC CC SITE 41 42 CLEAVAGE (BY ADENOVIRUS PROTEASE)
CC CC (POTENTIAL).
CC CC SEQUENCE 70 AA; 7769 MW; 015EA3C593558C49 CRC64;
CC CC
CC CC Query Match 63.0% Score 34; DB 1; Length 70;
CC CC Best Local Similarity 77.8%; Pred.No.7.2;
CC CC Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC CC
CC CC Qy 4 RYRVRRVRR 12
CC CC Db 27 RLVRRVARR 35
CC CC
CC CC RESULT 5
CC CC RL19_HUMAN STANDARD; PRT; 196 AA.
CC CC AC P14118; P22908;
CC CC DT 01-JAN-1990 (Rel. 13, Created)
CC CC DT 01-JAN-1990 (Rel. 13, Last sequence update)
CC CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC CC DE 60S ribosomal protein L19.
CC CC GN RPL19.
CC CC OS Homo sapiens (Human),
CC CC OS Mus musculus (Mouse), and
CC CC OS Rattus norvegicus (Rat).
CC CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CC CC CC NCBI TaxID=9606, 10090, 10116;
CC CC CC [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RC SPECIES=Human;
CC CC RA MEDLINE=92285147; PubMed=1598220;
CC CC RA Kumabe T., Schma Y., Yamamoto T.;
CC CC RT "Human cDNAs encoding elongation factor 1 gamma and the ribosomal
CC CC RT protein L19.";
CC CC RL Nucleic Acids Res. 20:2598-2598(1992).
CC CC CC [2]
CC CC RP SEQUENCE FROM N.A.
CC CC RC SPECIES=Human;
CC CC RA MEDLINE=93185086; PubMed=8095182;
CC CC RA Henry J.L., Coggin D.L., King C.R.;
CC CC RT "High-level expression of the ribosomal protein L19 in human breast
CC CC RT tumors that overexpress erbB-2.";
CC CC RL Cancer Res. 53:1403-1408(1993).
CC CC CC [3]
CC CC RP SEQUENCE FROM N.A.
CC CC RC SPECIES=Human; TISSUE=Muscle;
CC CC RA Strausberg R.;
CC CC RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
CC CC CC [4]
CC CC RP SEQUENCE FROM N.A.
CC CC RC SPECIES=Mouse;
CC CC RA MEDLINE=91090840; PubMed=1702292;
CC CC RA Nakamura T., Ono M., Marriage-Jamson R., Hillova J., Hill M.;
CC CC RT "Nucleotide sequence of mouse L19 ribosomal protein cDNA isolated in

```

screening with tire oncogene probes." ;
RT DNA Cell Biol. 6:697-703(1990).
[5]
SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aitawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamatake I.,
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischiemann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nkato I., Pesole G., Quackenbush J.,
RA Schirral L.M., Straubli F., Suzuki R., Tomita M., Wagner I., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Botelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guentrich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoedach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Teyo-Oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection." ;
RL Nature 403:665-690(2001).
[6]
SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RA Strauberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC SPECIES=Cat;
RX MEDLINE=87109220; PubMed=3542997;
RA Chan Y.-L., Lin A., McNally J., Peleg D., Meynhas O., Wool I.G.;
RT "The primary structure of rat ribosomal protein L19. A determination
from the sequence of nucleotides in a cDNA and from the sequence of
amino acids in the protein." ;
RL J. Biol. Chem. 262:1111-1115(1987).
[8]
SEQUENCE FROM N.A.
RC SPECIES=Cat;
RX MEDLINE=95309903; PubMed=7789970;
RA Davies B., Fried M.;
RT "The L19 ribosomal protein gene (RPL19): gene organization,
chromosomal mapping, and novel promoter region." ;
RL Genomics 25:372-380(1995).
-1- SIMILARITY: BELONGS TO THE L19E FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
EMBL; X63527; CAA45090.1; --
EMBL; S56985; AAB25672.1; --
EMBL; BC000530; AAH00530.1; --
EMBL; BC013016; AAH13016.1; --
EMBL; M62952; AAB48630.1; --
EMBL; AK010440; BAB26941.1; --
EMBL; BC010710; AAH10710.1; --
EMBL; J02650; AAA42071.1; --
EMBL; X82202; CAA57685.1; --
PIR; S09560; R5RT19.
PIR; A36554; A36554.
PIR; S22656; S22656.
PIR; A48992; A48992.
Genew; HGNC:10312; RPL19.
MTM; 180466; --

DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein RV1002C.
 GN RV1002C OR MTC1237.17C OR MT1031.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae;
 OC Actinomycetiales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=11773;
 RX MEDLINE=98295987; PubMed=9634230;
 RC STRAIN=H37Rv;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekela F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jaffe S., Kirogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Stalson J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.F., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 39.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z94752; CAB08157.1; -;
 DR EMBL; AE006986; AAK45281.1; -;
 DR TIGR; MT1031; -;
 DR Tuberculist; RV1002C; -;
 DR InterPro; IPR003342; PWT.
 DR Pfam; PF02366; PWT; 1.
 KW Hypothetical protein; Transferase; Glycosyltransferase; Transmembrane;
 KW Complete proteome.
 FT TRANSMEM 101 121 POTENTIAL.
 FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 153 173 POTENTIAL.
 FT TRANSMEM 207 229 POTENTIAL.
 FT TRANSMEM 263 283 POTENTIAL.
 FT TRANSMEM 400 420 POTENTIAL.
 FT TRANSMEM 424 444 POTENTIAL.
 FT TRANSMEM 460 480 POTENTIAL.
 SQ SEQUENCE 503 AA; 55531 MW; EB6226A8F1442C6 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 503;
 Best Local Similarity 70.0%; Pred. No. 27;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VVRVRRVVR 11
 DB 120 VVRVRRISR 129

RESULT 3
 URK_CLOPE STANDARD; PRT; 208 AA.
 ID URK_CLOPE
 AC O6X16;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
 DE monophosphokinase).
 GN UDK OR CPE1770.
 OS Clostridium perfringens.
 OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 NCBI_TaxID=1502;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / Type A;
 RC PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hatford M., Kohara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + UMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AP003191; BAB81476.1; -;
 DR InterPro; IPR001324; PRK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TRIGRAMS; TRGR00235; udk; 1.
 KW Transferase; Kinase; ATP-binding; Complete proteome.
 FT NP BIND 11 18 ATP (POTENTIAL).
 SQ SEQUENCE 208 AA; 23884 MW; 941BC368740DD5E CRC64;

Query Match 64.8%; Score 35; DB 1; Length 208;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVRVRRVVR 11
 DB 136 VVRVRRVVR 144

RESULT 4
 L2MU_ADEB2 STANDARD; PRT; 70 AA.
 ID L2MU_ADEB2
 AC Q96626;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Late L2 mu core protein precursor (pmu) (Protein X).
 GN PX.
 OS Bovine adenovirus type 2 (Mastadenovirus bo2).
 OS Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OC NCBI_TaxID=114429;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Rueval M., Harrach B., Banreivi A., Evans P., Benko M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE ROLE OF THE PRECURSOR MIGHT BE TO CONDENSE THE

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2003, 11:40:31 ; Search time 2.42553 Seconds

(without alignments)
205.199 Million cell updates/sec

Title: US-10-079-075-4

Sequence: 1 RVRRVRRVRR 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	68.5	105	RL31_AERPE	Q9YD25 aeropyrum p
2	36	66.7	503	YA02_MYCTU	O05586 mycobacteri
3	35	64.8	208	URK_ELORE	Q8XJ16 clostridium
4	34	63.0	70	L2MU_ADEB2	Q96626 bovine aden
5	34	63.0	196	RL19_HUMAN	P14118 homo sapien
6	34	63.0	211	URK_BACHD	Q9K488 bacillus ha
7	34	63.0	310	YFCEB_ECOLI	P39199 escherichia
8	34	63.0	310	YFCEB_ECOLI	P39199 escherichia
9	34	63.0	319	CBIB_SALTY	O05600 salmonella
10	34	63.0	613	YE40_ARCFU	O28832 archaeoglob
11	33	61.1	32	PRT_ORYLA	Q91185 oryzias lat
12	33	61.1	135	RM16_PROMI	Q91185 oryzias lat
13	33	61.1	314	YL18_VIBCH	Q9K683 vibrio chol
14	33	61.1	317	RPSD_STRAU	B27785 streptomyce
15	33	61.1	336	Y724_MERTH	O26820 methanobact
16	33	61.1	335	RFE_HAEIN	P45341 haemophilus
17	33	61.1	367	TROD_TREPA	P96119 treponema p
18	33	61.1	369	DNAJ_THEMA	O9WZ93 thermotoga
19	33	61.1	396	HRDA_STRCO	P18182 streptomyce
20	33	61.1	399	Y199_ARCFU	O28380 archaeoglob
21	33	61.1	710	RNR_THEMA	O9WZ11 thermotoga
22	33	61.1	847	ENV_HV1W2	P05880 human immun
23	33	61.1	848	ENV_HV1JR	P20871 human immun
24	33	61.1	853	ENV_HV1Z2	P12487 human immun
25	33	61.1	855	ENV_HV1Z6	P04580 human immun
26	33	61.1	856	ENV_HV1W1	P31872 human immun
27	33	61.1	868	ENV_HV1C4	P05879 human immun
28	33	59.3	88	RL31_SULTO	P54051 methanococc
29	32	59.3	92	R37A_METJA	O9YAD3 aeropyrum p
30	32	59.3	187	YK06_AERPE	O11078 caenorhabdi
31	32	59.3	187	YT61_CAPEL	P06696 staphylococ
32	32	59.3	361	TNPA_STRAM	O8XAE7 escherichia
33	32	59.3	367	WECA_ECO57	

34	32	59.3	367	1	WECA_ECOLI	P24235 escherichia
35	32	59.3	367	1	WECA_SALTY	Q8Z386 salmonella
36	32	59.3	367	1	WECA_SALTY	Q91677 salmonella
37	32	59.3	449	1	MES3_INPAC	O04719 autographa
38	32	59.3	484	1	DEF51_HUMAN	P57764 homo sapien
39	32	59.3	541	1	YD18_MYCTU	O10631 mycobacteri
40	32	59.3	582	1	SYP_MYCTU	O05814 mycobacteri
41	32	59.3	678	1	GSPD_AERHY	P31780 aeromonas h
42	32	59.3	678	1	GSPD_AERSA	P45778 aeromonas s
43	32	59.3	852	1	ENV_HV1S3	P19549 human immun
44	32	59.3	877	1	SYA_SYNY3	P74423 synchocyst
45	32	59.3	880	1	SYA_ANASP	Q9YD4 anaerobac sp

ALIGNMENTS

RESULT 1	ID	RL31_AERPE	STANDARD;	PRT;	105 AA.
AC	Q9YD25;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	50S ribosomal protein L31e.				
GN	RL31E OR APE1087.				
OS	Aeropyrum pernix.				
OC	Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;				
OC	Desulfurococcaceae; Aeropyrum.				
OX	NCBI_TaxID=56636;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=K1;				
RC	MEDLINE=99310339; PubMed=10382966;				
RA	Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,				
RA	Jin-no K., Takahashi M., Sekine M., Baba S.-T., Anai A., Kosugi H.,				
RA	Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,				
RA	Takamitsu M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,				
RA	Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,				
RA	Nakamura Y., Nomura N., Sako Y., Kikuchi H.,				
RT	"Complete genome sequence of an aerobic hyper-thermophilic				
RT	crenarchaeon, Aeropyrum pernix K1.";				
RT	DNA Res. 6:83-101(1999).				
CC	-1- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.				
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	CC between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	CC the European Bioinformatics Institute. There are no restrictions on its				
CC	CC use by non-profit institutions as long as its content is in no way				
CC	CC modified and this statement is not removed. Usage by and for commercial				
CC	CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/				
CC	CC or send an email to license@isb-sib.ch).				
CC	CC				
DR	EMBL; AB000060; BAA80072.1; -				
DR	InterPro; IPR000054; Ribosomal_L31e.				
DR	Pfam; PF01198; Ribosomal_L31e; 1.				
DR	PROSITE; PS01144; RIBOSOMAL_L31E; FALSE_NEG.				
KW	Ribosomal protein; Complete proteome.				
SO	SEQUENCE 105 AA; 12527 MW; 75EDF799E74A098 CR664;				
Query Match	68.5%; Score 37; DB 1; Length 105;				
Best Local Similarity	66.7%; Pred. No. 3.3;				
Matches	8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;				
QY	1 RVRRVRRVRR 12				
DB	28 RAVRVRRVRR 39				
RESULT 2	YA02_MYCTU				
ID	YA02_MYCTU	STANDARD;	PRT;	503 AA.	
AC	O05586;				